

FIGURE 1

2/75

FIGURE 2

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA22779
><subunit 1 of 1, 493 aa, 1 stop
><MW: 57104, pI: 7.67, NX(S/T): 2
MRPLCVTCWWLGLLAAMGAVAGQEDGFEGTEEGSPREFIYLNRYKRAGESQDKCTYTFIVPQ
QRVTGAICVNSKEPEVLLNENRVHVKQELELLNNELLKQKRQIETLQQLVEVDGGIVSEVKLLR
KESRNMNSRVTQLYMQLLHEIIRKRDNALELSQLENRILNQADMLQLASKYKDLEHKYQHL
ATLAHNQSEIIIAQLEEHQCRVPSARVPVQPPPAAPPRVYQPPTYNRIINQISTNEIQSDQNL
KVLPPPLPTMPTLTSLPSSTDKPSGPWRDCLQALEDGHDTSIIYLVKPENTNRLMQVWCDQR
HDPGGWTVIQRRLDGSVNFFRNWETKQGGFNI DGEYWLGLENIYWLTNQGNKLLVTMEDW
SGRKVFAEYASFRLEPESEYYKLRLGRYHGNAGDSFTWHNGKQFTTLDRDHDVYTGNCAYHQ
KGGWWYNACAHSNLNGVWYRGGHYRSRYQDGVYWAEFRGGSYSLLKKVMMIRPNPNTFH
```

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 164-168, 192-196

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 124-128

Tyrosine kinase phosphorylation sites.

amino acids 177-184, 385-393, 385-394, 461-468

N-myristoylation sites.

amino acids 12-18, 18-24, 22-28, 29-35, 114-120, 341-347,
465-471, 473-479

Amidation site.

amino acids 373-377

Fibrinogen beta and gamma chains C-terminal domain signature.

amino acids 438-451

Fibrinogen beta and gamma chains C-terminal domain proteins.

amino acids 305-343, 365-402, 411-424, 428-458

Trehalase proteins.

amino acids 275-292

CCACACGCGTCCGGCGCCGTGGCCTCGCGTCCATCTTTGCCGTTCTCTCGGACCTGTCCACAAA
GGAGTCGCGCCGCCGCCGCCGCCCTCCCTCCGGTGGGCCCCGGGAGGTAGAGAAAGTCAGT
GCCACAGCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCTGAGGGT
TGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACCGGACTGGG
GCCGGGGTAGGCTCTGGAAAGGGCCCCGGGAGAGAGGTGGCGTTGGTCAGAACCTGAGAAACA
GCCGAGAGGTTTTCCACCGAGGCCCGCGCTTGAGGGATCTGAAGAGGTTCTTAGAAGAGGGT
GTTCCCTCTTTGCGGGGTCCTCACCAGAAGAGGTTCTTGGGGGTCGCCCTTCTGAGGAGGCT
GCGGCTAACAGGGCCCAGAACTGCCATTGGATGTCCAGAATCCCCTGTAGTTGATAATGTTG
GGAATAAGCTCTGCAACTTTCTTTGGCATTGAGTTGTTAAAAACAAATAGGATGCAATTCC
TCAACTCCAGGTTATGAAAACAGTACTTGGAACCTGAAACCTACCTAAATGATCGTCTTTG
GTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGGCCAGGGTCTGTTGTTGACTCTCGAAGAG
CACATAGCCCACTTCCTAGGGACTGGAGGTGCCGCTACTACCATGGGTAATTCCTGTATCTG
CCGAGATGACAGTGGAACAGATGACAGTGTTGACACCCAACAGCAACAGGCCGAGAACAGTG
CAGTACCCACTGCTGACACAAGGAGCCAACCACGGGACCCTGTTGCGGCCACCAAGGAGGGGC
CGAGGACCTCATGAGCCAAGGAGAAAGAAACAAAATGTGGATGGGCTAGTGTTGGACACACT
GGCAGTAATACGGACTCTTGTTAGATAAGTAAGTATCTGACTCACGGTCACCTCCAGTGGAAT
GAAAAGTGTTCTGCCCGGAACCATGACTTTAGGACTCCTTCAGTTCCTTTAGGACATACTCG
CCAAGCCTTGTTGCTCACAGGGCAAAGGAGAATATTTTAATGCTCCGCTGATGGCAGAGTAA
TGATAAGATTTGATGTTTTTGCTTGCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTC
TGTAGCAGAAAACACGATAAAGCTATGATCTTTATTAGAG

4/75

FIGURE 4

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26846

<subunit 1 of 1, 117 aa, 1 stop

<MW: 12692, pI: 7.50, NX(S/T): 0

MIVFGWAVFLASRLGQGLLLTLEEHIAHFLGTGGAATTMGNSCICRDDSGTDDSVDTQQQQ

AENSAVPTADTRSQPRDPVRPPRRGRGPHEPRRKKQNV DGLVLDTLAVIRTLVDK

Important features:

Signal peptide:

amino acids 1-16

N-myristoylation sites.

amino acids 18-24, 32-38, 34-40, 35-41, 51-57

FIGURE 5

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG
CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGC
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGGAGACAGCGGACAAGCAG
CGGAGGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCAGCAAAAGAAGCGGTGGTGGTGGG
CGTCGTGGCC**ATG**GCGGCGGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTGACAGCCCCAGCAAAGGCAAGACCAGCTGC
GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT
CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACCTTTTCACACCTGAGTGCA
AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC
AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC
G**TAG**CCAGTGAGGGCAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTGAGTGACATTTACCAAACAAACAGG
CAGAGTTCACCTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

6/75

FIGURE 6

```

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498
><subunit 1 of 1, 245 aa, 1 stop
><MW: 27564, pI: 10.18, NX(S/T): 1
MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGTSCDKNKLNVF SRVKLFGSKKRRRRRP
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLI PVGLRVVAIQGVQTKLYLA
MNSEGYYLTSELF TPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMHNEST
  
```

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

7175

FIGURE 8

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503
><subunit 1 of 1, 247 aa, 1 stop
><MW: 27702, pI: 10.36, NX(S/T): 2
MAAAIASGLIRQKRQAREQHWD RPSASRRRSSPSKNRGLCNGNLVDIFSKVRI FGLKKRRLR
RQDPQLKGIVTRLYCRQGYLQMHPDGALDGTKDDSTNSTLFNLI PVGLRVVAIQGVKTGLY
IAMNGEGYLYPSELF TPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETV PKPGVTPSKSTSASAIMNGGKPVNKS KTT
```

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

9/75

10/75

FIGURE 10

MPGIKRILTVTILALCLPSPGNAQAQCTNGFDLDRQSGQCLDIDECRTIPEACRGDMMCVNQ
NGGYLCIPRTNPVYRGPYSNPYSTPYSGPYPAAPPLSAPNYPTISRPLICRFGYQMDESNO
CVDVDECATDSHQCNPTQICINTEGGYTCSCTDGYWLLEGQCLDIDECRYGYCQQLCANVPG
SYSCTCNPGFTLNEDGRSCQDVNECATENPCVQTCVNTYGS LICRCDPGYELEEDGVHCSDM
DECSFSEFLCQHECVNQPGTYFCSCPPGYILLDDNRSCQDINECEHRNHTCNLQQTCYNLQG
GFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPF TILYRDMDVVSGRSVPADIFQMQATT
RYPGAYYIFQIKSGNEGREFYMRQTGPISATLVMTRPIKGPRIQLDLEMITVNTVINFRGS
SVIRLRIVVSQYPF

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 283-287, 296-300

N-myristoylation sites.

amino acids 21-27, 64-70, 149-155, 186-192, 226-232, 242-248,
267-273, 310-316

Aspartic acid and asparagine hydroxylation sites.

amino acids 144-156, 181-193, 262-274

Cell attachment sequence.

amino acids 54-57

Calcium-binding EGF-like.

amino acids 131-166, 172-205, 211-245, 251-286

FIGURE 11

CAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTC
GACCTCGACCCACGCGTCCGAACACAGGTCCTTGTTGCTGCAGAGAAGCAGTTGTTTTGCTG
GAAGGAGGGAGTGCGCGGGCTGCCCGGGCTCCTCCCTGCCGCCTCCTCTCAGTGATGGTT
CCAGGCACCCTGTCTGGGGCAGGGAGGGCACAGGCCTGCACATCGAAGGTGGGGTGGGACCA
GGCTGCCCCCTCGCCCCAGCATCCAAGTCCCTTGGGCGCCCCGTGGCCCTGCAGACTCTCA
GGGCTAAGGTCTCTGTTGCTTTTTGGTTCCACCTTAGAAGAGGCTCCGCTTGACTAAGAGT
AGCTTGAAGGAGGCACCA**ATG**CAGGAGCTGCATCTGCTCTGGTGGGCGCTTCTCCTGGGCCTG
GCTCAGGCCTGCCCTGAGCCCTGCGACTGTGGGGAAAAGTATGGCTTCCAGATCGCCGACTG
TGCCCTACCGCGACCTAGAATCCGTGCCGCCTGGCTTCCCGGCCAATGTGACTACACTGAGCC
TGTGAGCCAACCGGCTGCCAGGCTTGCCGGAGGGTGCCTTCAGGGAGGTGCCCTGCTGCAG
TCGCTGTGGCTGGCACACAATGAGATCCGCACGGTGGCCGCCGGAGCCCTGGCCTCTCTGAG
CCATCTCAAGAGCCTGGACCTCAGCCACAATCTCATCTCTGACTTTGCCTGGAGCGACCTGC
ACAACCTCAGTGCCCTCCAATTGCTCAAGATGGACAGCAACGAGCTGACCTTCATCCCCCGC
GACGCCTTCCGCAGCCTCCGTGCTCTGCGCTCGCTGCAACTCAACCACAACCGCTTGACAC
ATTGGCCGAGGGCACCTTCACCCCGCTCACCGCGCTGTCCCACCTGCAGATCAACGAGAACC
CCTTCGACTGCACCTGCGGCATCGTGTGGCTCAAGACATGGGCCCTGACCACGGCCGTGTCC
ATCCCGGAGCAGGACAACATCGCCTGCACCTCACCCCATGTGCTCAAGGGTACACCGCTGAG
CCGCCTGCCGCCACTGCCATGCTCGGCGCCCTCAGTGCAGCTCAGCTACCAACCCAGCCAGG
ATGGTGCCGAGCTGCGGCCTGGTTTTGTGCTGGCACTGCACTGTGATGTGGACGGGCAGCCG
GCCCCCTCAGCTTCACTGGCACATCCAGATAACCCAGTGGCATTGTGGAGATCACCAGCCCCAA
CGTGGGCACTGATGGGCGTGCCCTGCCTGGCACCCCTGTGGCCAGCTCCCAGCCGCGCTTCC
AGGCCTTTGCCAATGGCAGCCTGCTTATCCCCGACTTTGGCAAGCTGGAGGAAGGCACCTAC
AGCTGCCTGGCCACCAATGAGCTGGGCAGTGTGAGAGCTCAGTGGACGTGGCACTGGCCAC
GCCCCGTGAGGGTGGTGAGGACACACTGGGGCGCAGGTTCCATGGCAAAGCGGTTGAGGGAA
AGGGCTGCTATACGGTTGACAACGAGGTGCAGCCATCAGGGCCGGAGGACAATGTGGTTCATC
ATCTACCTCAGCCGTGCTGGGAACCCTGAGGCTGCAGTCGCAGAAGGGGTCCCTGGGCAGCT
GCCCCCAGGCCTGCTCCTGCTGGGCCAAAGCCTCCTCCTCTTCTTCTTCTCCTCACCTCCTTCT**T**
AGCCCCACCCAGGGCTTCCCTAACTCCTCCCCTTGCCCCCTACCAATGCCCTTTTAAGTGCTG
CAGGGGTCTGGGGTTGGCAACTCCTGAGGCCTGCATGGGTGACTTCACATTTTCTTACCTCT
CCTTCTAATCTCTTCTAGAGCACCTGCTATCCCCAACTTCTAGACCTGCTCCAAACTAGTGA
CTAGGATAGAATTTGATCCCCTAACCTCACTGTCTGCGGTGCTCATTGCTGCTAACAGCATTG
CCTGTGCTCTCCTCTCAGGGGCAGCATGCTAACGGGGCGACGTCCTAATCCAACCTGGGAGAA
GCCTCAGTGGTGGAAATTCAGGCACTGTGACTGTCAAGCTGGCAAGGGCCAGGATTGGGGGA
ATGGAGCTGGGGCTTAGCTGGGAGGTGGTCTGAAGCAGACAGGGAATGGGAGAGGAGGATGG
GAAGTAGACAGTGGCTGGTATGGCTCTGAGGCTCCCTGGGGCCTGCTCAAGCTCCTCCTGCT
CCTTGCTGTTTTCTGATGATTTGGGGGCTTGGGAGTCCCTTTGTCTCATCTGAGACTGAAA
TGTGGGGATCCAGGATGGCCTTCCTTCTTACCCTTCTCCCTCAGCCTGCAACCTCTAT
CCTGGAACCTGTCTCCCTTTCTCCCCAACTATGCATCTGTTGTCTGCTCCTCTGCAAAGGC
CAGCCAGCTTGGGAGCAGCAGAGAAATAAACAGCATTTCTGATGCCAAAAAAAAAAAAAAAAA
AAGGGCGGCCGCGACTCTAGAGTCGACCT

FIGURE 12

MQELHLLWWALLLGLAQACPEPCDCGEKYGFQIADCAYRDLESVPPGFANVTTLSSLNRL
PGLPEGAFREVPLLQSLWLAHNEIRTVAAGALASLSHLKSLDLSHNLISDFAWSDLHNLSAL
QLLKMSNELTFIPRDAFRSLRALRSLQLNHNRLHTLAEGTFTPLTALSHLQINENPFDCCTC
GIVWLKTWALTAVSIPEDNIACTSPHVLKGTPLSRLPPLPCSAPSVQLSYQPSQDGAELR
PGFVLALHCDVDGQPAPQLHWHIQIPSGIVEITSPNVGTDGRALPGTPVASSQPRFQAFANG
SLLIPDFGKLEEGTYSCLATNELGSAESSVDVALATPGEGGEDTLGRRFHGKAVEGKGCYTV
DNEVQPSGPEDNVVIIYLSRAGNPEAAVAEGVPGQLPPGLLLLQSLLLFFFLTSF

Important features of the protein:

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 403-418

N-glycosylation sites.

amino acids 51-55, 120-124, 309-313

Tyrosine kinase phosphorylation site.

amino acids 319-326

N-myristoylation sites.

amino acids 14-20, 64-70, 92-98, 218-224, 294-300, 323-329,
334-340, 350-356, 394-400

Amidation site.

amino acids 355-359

Leucine rich repeats.

amino acids 51-74, 75-98, 99-122, 123-146, 147-170

Leucine rich repeat C-terminal domain.

amino acids 180-230

CCAGGCCGAGGAGGCGACGCGGCCAGCGCTCTAAACGGGAACAGCCCTGGCTTGAGGGAGCTGC
AGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCGGCACGAGGAGTTTTCCCGGC
AGCGAGGAGGTCCTGAGCAGC**ATG**CCCCGGAGGAGCGCCTTCCCTGCCGCCGCGCTCTGGCT
CTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGCCGGGCCGCCGAGGAGGAGA
GCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCATAGGATTTGAAGAAGATATC
CTGATTGTTTTAGAGGGGAAAATGGCACCTTTTACACATGATTTTCAAAAAGCGCAACAGAG
AATGCCAGCTATTCTGTCAATATCCATTCCATGAATTTTACCTGGCAAGCTGCAGGGCAGG
CAGAATACTTCTATGAATTCCTGTCCTTGCGCTCCCTGGATAAAGGCATCATGGCAGATCCA
ACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCATCAGTTGTTCAAGTTGGTTT
CCCATGTCTTGGAACACAGGATGGGGTGGCAGCATTTGAAGTGGATGTGATTGTTATGAATT
CTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCTTCTTTAAACATGTCAACAA
GCTGAGTGCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAATGAAAGACGCATCTGCGAGTG
TCCTGATGGGTTCACGGACCTCACTGTGAGAAAGCCCTTTGTACCCACGATGTATGAATG
GTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTGGATTCTATGGAGTGAAGTGT
GACAAAGCAAACCTGCTCAACCACCTGCTTTAATGGAGGGACCTGTTTCTACCCTGGAAAATG
TATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAGCAAATGCCCAACCCCTGTC
GAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTTCCAAAGGTTACCAGGGAGAC
CTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACATGGAACCTGCCATGAACCCAA
CAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAATAAAAGGTACGAAGCCAGCC
TCATACATGCCCTGAGGCCAGCAGGCGCCCAGCTCAGGCAGCACACGCCTTCACTTAAAAAG
GCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG**TGA**ACTCCGACATCTGAAAC
GTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCATGTGTTGAATGTTCAAATAA
TGTTTCATTACACTTAAGAATACTGGCCTGAATTTTATTAGCTTCATTATAAATCACTGAGCT
GATATTTACTCTTCCCTTTTAAGTTTTCTAAGTACGTCTGTAGCATGATGGTATAGATTTTCT
TGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGATCAGGTTAAATTTTCAGTG
TGTAGTTGGCAGATATTTTCAAATTTACAATGCATTTATGGTGTCTGGGGGCAGGGGAACAT
CAGAAAGGTTAAATTTGGGCAAAAATGCGTAAGTCACAAGAATTTGGATGGTGCAGTTAATGT
TGAAGTTACAGCATTTTCAGATTTTATTGTCAGATATTTAGATGTTTGTTACATTTTTTAAAAA
TTGCTCTTAATTTTTTAACTCTCAATACAATATATTTTGACCTTACCATTATTCCAGAGATT
CAGTATTAATAAAAAAAAAAATTACACTGTGGTAGTGGCATTTAAACAATATAATATATTCTA
AACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCATTGGCTTGAAGCAATATAATA
TATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTATACTGTTTGTATGTATAAAAT
AAAGGTGCTGCTTTAGTTTTTTTGGAAAAA

FIGURE 14

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESELYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMIPAIPVNIHSMNFTWQAAGQAEIFYEFLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPPGGC
RNGGFCNERRICECPDGFHGHPCHEKALCTPRCMNGGLCVTPGFCICPPPGFYGVNCDKANCST
TCFNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAERRDP
PESNYIW

Signal sequence.

amino acids 1-28

N-glycosylation sites.

amino acids 88-92, 245-249

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

FIGURE 15

AGAACCTCAGAAATGTGAGTTATTTGGGAATGGCTGTTTGTAATGTCCTTACGTAAGCCAA
GAGGAGGTCTTGACTTGGGGTCCCAGGGGTACCGCAGATCCCAGGGACTGGAGCAGCACTAG
CAAGCTCTGGAGGATGAGCCAGGAGTCTGGAATTGAGGCTGAGCCAAAGACCCCAGGGCCGT
CTCAGTCTCATAAAAGGGGATCAGGCAGGAGGAGTTTGGGAGAAACCTGAGAAGGGCCTGAT
TTGCAGCATC**ATG**ATGGGCCTCTCCTTGGCCTCTGCTGTGCTCCTGGCCTCCCTCCTGAGTC
TCCACCTTGGAAGTGGCACACGTGGGAGTGACATATCCAAGACCTGCTGCTTCCAATACAGC
CACAAGCCCCCTTCCCTGGACCTGGGTGCGAAGCTATGAATTCACCAGTAACAGCTGCTCCCA
GCGGGCTGTGATATTCACTACCAAAGAGGCAAGAAAGTCTGTACCCATCCAAGGAAAAAAT
GGGTGCAAAAATACATTTCTTTACTGAAACTCCGAAACAATTG**TGA**CTCAGCTGAATTTTC
ATCCGAGGACGCTTGGACCCCGCTCTTGGCTCTGCAGCCCTCTGGGGAGCCTGCGGAATCTT
TTCTGAAGGCTACATGGACCCGCTGGGGAGGAGAGGGTGTTTCCTCCCAGAGTTACTTTAAT
AAAGGTTGTTCATAGAGTTGAAA
AAA

FIGURE 16

MMGLSLASAVLLASLLSLHLGTATRGSDISKTCCEFQYSHKPLPWTWVRSYEFTSNSCSQRAV
IFTTKRGKKVCTHPRKKWVQKYISLLKTPKQL

Important features of the protein:

Signal peptide:

amino acids 1-23

N-myristoylation sites.

amino acids 3-9, 26-32

Amidation site.

amino acids 68-72

Small cytokines (intecrine/chemokine).

amino acids 23-88

FIGURE 17

GCGAGAACCTTTGCACGCGCACAAACTACGGGGACGATTTCTGATTGATTTTGGCGCTTCGATCCACCCTCC
TCCCTTCTCATGGGGACTTTGGGGACAAAGCGTCCCGACCGCCTCGAGCGCTCGAGCAGGGCGCTATCCAGGAGC
CAGGACAGCGTCGGGAACCAGACCATGGCTCCTGGACCCCAAGATCCTTAAGTTTCGTCGTCCTTCATCGTCGCGG
TTCTGCTGCCGGTCCGGGTTGACTCTGCCACCATCCCCGGCAGGACGAAGTTCCCCAGCAGACAGTGGCCCCA
CAGCAACAGAGGGCGCAGCCTCAAGGAGGAGGAGTGTCCAGCAGGATCTCATAGATCAGAATATACTGGAGCCTG
TAACCCGTGCACAGAGGGTGTGGATTACACCATTGCTTCCAACAATTTGCCTTCTTGCTGCTATGTACAGTTT
GTAATCAGGTCAAACAAATAAAAGTTTCTGTACCACGACCAGAGACACCGTGTGTGTCAGTGTGAAAAAGGAAGC
TTCCAGGATAAAAACTCCCCTGAGATGTGCCGGACGTGTAGAACAGGGTGTCCAGAGGGATGGTCAAGGTCAG
TAATTGTACGCCCCGGAGTGACATCAAGTGCAAAAATGAATCAGCTGCCAGTTCCACTGGGAAAACCCAGCAG
CGGAGGAGACAGTGACCACCATCCTGGGGATGCTTGCCCTCTCCCTATCACTACCTTATCATCATAGTGGTTTTA
GTCATCATTTTAGCTGTGGTTGTGGTTTTCATGTCCGAAGAATTCATTTCTTACCTCAAAGGCATCTG
CTCAGGTGGTGGAGGAGGTCCCGAACGTGTGCACAGAGTCCCTTTTCCGGCGGCGTTCATGTCCCTTCACGAGTTT
CTGGGGCGGAGGACAATGCCCGCAACGAGACCCTGAGTAACAGATACTTGCAGCCACCCAGGTCTCTGAGCAG
GAAATCCAAGGTGAGGAGCTGGCAGAGCTAACAGGTGTGACTGTAGAGTCGCCAGAGGAGCCACAGCGTCTGCT
GGAACAGGCAGAAGCTGAAGGGTGTGAGAGGAGGAGGCTGCTGGTTCCAGTGAATGACGCTGACTCCGCTGACA
TCAGCACCTTGCTGGATGCCTCGGCAACACTGGAAGAAGGACATGCAAGGAAACAATTCAGGACCAACTGGTG
GGCTCCGAAAAGCTCTTTTATGAAGAAGATGAGGCAGGCTCTGCTACGTCTGCTGCCTGCTGAAGAAATCTCTCAG
GAAACCAGAGCTTCCCTCATTTACCTTTTCTCCTACAAAGGGAAGCAGCCTGGAAGAAAACAGTCCAGTACTTGA
CCCATGCCCAACAACTCTACTATCCAATATGGGGCAGCTTACCAATGGTCCCTAGAATTTGTTAACGCACTT
GGAGTAATTTTTATGAATACTGCGTGTGATAAGCAAACGGGAGAAATTTATATCAGATTCTTGGCTGCATAGT
TATACGATTGTGTATTAAGGGTCGTTTTAGGCCACATGCGGTGGCTCATGCCTGTAATCCCAGCACTTTGATAG
GCTGAGGCAGGTGGATTGCTTGAGCTCGGGAGTTTGAGACCAGCCTCATCAACACAGTGAAACTCCATCTCAAT
TTAAAAAGAAAAAAGTGGTTTTAGGATGTCATTCTTTGCAGTTCTTCATCATGAGACAAGTCTTTTTTCTGC
TTCTTATATTGCAAGCTCCATCTCTACTGGTGTGTGCATTTAATGACATCTAACTACAGATGCCGCACAGCCAC
AATGCTTTGCCTTATAGTTTTTAACTTTAGAACGGGATTATCTTGTTATTACCTGTATTTTCAGTTTCGGATA
TTTTTGACTTAATGATGAGATTATCAAGACGTAGCCCTATGCTAAGTCATGAGCATATGGACTTACGAGGGTTC
GACTTAGAGTTTTGAGCTTTAAGATAGGATTATTGGGGCTTACCCCACTTAATTAGAGAAAACATTTATTTG
CTTACTACTGTAGGCTGTACATCTCTTTTCCGATTTTGTATAATGATGTAAACATGGAAAAACCTTAGGAAAT
GCACTTATTAGGCTGTTTACATGGGTTGCCTGGATACAAATCAGCAGTCAAAAATGACTAAAAATATAACTAGT
GACGGAGGGAGAAATCCTCCCTCTGTGGGAGGCACTTACTGCATTCCAGTTCTCCCTCCTGCGCCCTGAGACTG
GACCAGGGTTTGATGGCTGGCAGCTTCTCAAGGGGCAGCTTGTCTTACTTGTAAATTTTAGAGGTATATAGCCA
TATTTATTTATAAATAAATATTTATTTATTTATTTATAAGTAGATGTTTACATATGCCAGGATTTTGAAGAGC
CTGGTATCTTTGGGAAGCCATGTGTCTGGTTTGTGCTGGGACAGTCATGGGACTGCATCTTCCGACTTGTC
CACAGCAGATGAGGACAGTGAGAATTAAGTTAGATCCGAGACTGCGAAGAGCTTCTCTTCAAGCGCCATTACA
GTTGAACGTTAGTGAATCTTGAGCCTCATTTGGGCTCAGGGCAGAGCAGGTGTTTATCTGCCCGGCATCTGCC
ATGGCATCAAGAGGGAAGAGTGGACGGTGTCTGGGAATGGTGTGAAATGGTTGCCGACTCAGGCATGGATGGGC
CCCTCTCGCTTCTGGTGGTCTGTGAAGTGGTCCCTGGGATGCCTTTTAGGGCAGAGATTCTTGAGCTGCGTTT
TAGGGTACAGATTCCCTGTTTGAGGAGCTTGGCCCCCTGTGAAGCATCTGACTCATCTCAGAGATATCAATTCT
TAAACACTGTGACAACGGGATCTAAAATGGCTGACACATTTGTCTTGTGTACGTTCCATTATTTTATTTAAA
AACCTCAGTAATCGTTTTAGCTTCTTTCCAGCAAACCTCTTCTCCACAGTAGCCAGTCTGGTAGGATAAATTA
CGGATATAGTCATTCTAGGGGTTTCAGTCTTTTCCATCTCAAGGCATTGTGTGTTTTGTTCGGGACTGGTTTG
GCTGGGACAAAGTTAGAACTGCCTGAAGTTCGCACATTAGATTGTTGTGTCCATGGAGTTTAGGAGGGGATG
GCCTTTCCGCTCTTCGCACTTCCATCCTCTCCCACTTCCATCTGGCGTCCCACACCTTGTCCCCTGCATTTCTG
GATGACACAGGGTGTGCTGCTGCTCCTAGTCTTTGCCCTTGTGCTGGGCTTCTGTGCAGGAGACTTGGTCTCAAAG
CTCAGAGAGAGCCAGTCCGGTCCCAGCTCCTTTGTCCCTTCCCTCAGAGGCCTTCCCTGAAGATGCATCTAGACT
ACCAGCCTTATCAGTGTTTAAGCTTATTCCTTTAACATAAGCTTCCCTGACAACATGAAATTGTTGGGGTTTTTT
GGCGTTGGTTGATTTGTTTAGGTTTTGCTTTATACCCGGGCCAAATAGCACATAACACCTGGTTATATATGAAA
TACTCATATGTTTATGACCAAAATAAATATGAAACCTCATRTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 18

MGLWGQSVPTASSARAGRYPGARTASGTRPWLLDPKILKFVVFIVAVLLPVRVDSATIPRQD
EVPQQTVAPQQRRSLKEEECPAGSHRSEYTGACNPCTEGVDYTIASNNLPSCLLCTVCKSG
QTNKSSCTTTTRDTCQCEKGSFQDKNSPEMCRTCRTGCPRGMVKVSNCTPRSDIKCKNESAA
SSTGKTPAAEETVTTILGMLASPHYHYLIIIIVVLVIIILAVVVVGFSCRKKFISYLGKICSGGG
GGPERVHRVLFRRRSCPSRVPGAEDNARNETLSNRYLQPTQVSEQEIQGQELAELTGVTVES
PEEPQRLLEQAEAEGCQRRRLVLPVNDADSADISTLLDASATLEEGHAKETIQDQLVGSEKL
FYEEDEAGSATSC

Important features of the protein:

Transmembrane domains:

amino acids 35-52, 208-230

N-glycosylation sites.

amino acids 127-131, 182-186, 277-281

Glycosaminoglycan attachment site.

amino acids 245-249

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 260-264

N-myristoylation sites.

amino acids 21-27, 86-92, 102-108, 161-167, 242-248, 270-276,
297-303, 380-386

ATP/GTP-binding site motif A (P-loop).

amino acids 185-193

TNFR/NGFR cysteine-rich region.

amino acids 99-139

FIGURE 19

GCGGCACCTGGAAGATGCGCCCATTTGGCTGGTGGCCTGCTCAAGGTGGTGTTCGTGGTCTTC
GCCTCCTTGTGTGCCTGGTATTCGGGGTACCTGCTCGCAGAGCTCATTCCAGATGCACCCCT
GTCCAGTGCTGCCTATAGCATCCGCAGCATCGGGGAGAGGCCTGTCCTCAAAGCTCCAGTCC
CCAAAAGGCAGAAAATGTGACCACTGGACTCCCTGCCCATCTGACACCTATGCCTACAGGTTA
CTCAGCGGAGGTGGCAGAAGCAAGTACGCCAAAATCTGCTTTGAGGATAACCTACTTATGGG
AGAACAGCTGGGAAATGTTGCCAGAGGAATAAACATTGCCATTGTCAACTATGTAACCTGGGA
ATGTGACAGCAACACGATGTTTTGATATGTATGAAGGCGATAACTCTGGACCGATGACAAAG
TTTATTCAGAGTGCTGCTCCAAAATCCCTGCTCTTCATGGTGACCTATGACGACGGAAGCAC
AAGACTGAATAACGATGCCAAGAATGCCATAGAAGCACTTGGAAGTAAAGAAATCAGGAACA
TGAAATTCAGGTCTAGCTGGGTATTTATTGCAGCAAAAGGCTTGGAACCTCCCTTCCGAAATT
CAGAGAGAAAAGATCAACCACTCTGATGCTAAGAACAACAGATATTCTGGCTGGCCTGCAGA
GATCCAGATAGAAGGCTGCATACCCAAAGAACGAAGCTGAACCTGCAGGGTCCTGAGTAAAT
GTGTTCTGTATAAACAAATGCAGCTGGAATCGCTCAAGAATCTTATTTTTCTAAATCCAACA
GCCCATATTTGATGAGTATTTTGGGTTTGTGTAAACCAATGAACATTTGCTAGTTGTATCA
AATCTTGGTACGCAGTATTTTATACCAGTATTTTATGTAGTGAAGATGTCAATTAGCAGGA
AACTAAAATGAATGGAAATTCTTAAAAAAAAA

FIGURE 20

MRPLAGGLLKVVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERPVLKAPVPKRQK
CDHWTPCPSDTYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGNVTAT
RCFDMYEGDNSGPMTKFIQSAAPKSLLFMVTYDDGSTRLNNDKNAIEALGSKEIRNMKFRS
SWVFIAAKGLELPSEIQREKINHSDAKNNRYSGWPAEIQIEGCI PKERS

Important features of the protein:

Signal peptide:

amino acids 1-20

N-glycosylation sites.

amino acids 120-124, 208-212

Glycosaminoglycan attachment site.

amino acids 80-84

N-myristoylation sites.

amino acids 81-87, 108-114, 119-125

FIGURE 21

CCGGGGAGGGGAGGGCCCGTCCCGCCCCCTCCCCGTCTCTCCCCGCCCCCTCCCCGTCCCTCCC
GCCGAAGCTCCGTCCCGCCCCGCGGGCCGGCTCCGCCCTCACCTCCCGGCCGCGGGCTGCCCTC
TGCCCGGGTTGTCCAAG**ATG**GAGGGCGCTCCACCGGGGTCGCTCGCCCTCCGGCTCCTGCTG
TTCGTGGCGCTACCCGCCTCCGGCTGGCTGACGACGGGCGCCCCCGAGCCGCCGCCGCTGTC
CGGAGCCCCACAGGACGGCATCAGAATTAATGTAACCTACACTGAAAGATGATGGGGACATAT
CTAAACAGCAGGTTGTTCTTAACATAACCTATGAGAGTGGACAGGTGTATGTAAATGACTTA
CCTGTAAATAGTGGTGTAAACCCGAATAAGCTGTCAGACTTTGATAGTGAAGAATGAAAATCT
TGAAAATTTGGAGGAAAAAGAATATTTTGGAAATTGTCAGTGTAAAGGATTTTAGTTCATGAGT
GGCCTATGACATCTGGTTCCAGTTTGCAACTAATTGTCATTCAAGAAGAGGTAGTAGAGATT
GATGGAAAACAAGTTCAGCAAAAGGATGTCACTGAAATTGATATTTTAGTTAAGAACCGGGG
AGTACTCAGACATTCAAACCTATACCCTCCCTTTGGAAGAAAGCATGCTCTACTCTATTTCTC
GAGACAGTGACATTTTATTTACCCTTCCTAACCTCTCCAAAAAGAAAGTGTTAGTTCACTG
CAAACCACTAGCCAGTATCTTATCAGGAATGTGGAAACCACTGTAGATGAAGATGTTTTACC
TGGCAAGTTACCTGAAACTCCTCTCAGAGCAGAGCCGCCATCTTCATATAAGGTAATGTGTC
AGTGGATGGAAAAGTTTAGAAAAGATCTGTGTAGGTTCTGGAGCAACGTTTTCCAGTATTC
TTTCAGTTTTTTGAACATCATGGTGGTTGGAATTACAGGAGCAGCTGTGGTAATAACCATCTT
AAAGGTGTTTTTCCCAGTTTCTGAATACAAAGGAATTCTTCAGTTGGATAAAGTGGACGTCA
TACCTGTGACAGCTATCAACTTATATCCAGATGGTCCAGAGAAAAGAGCTGAAAACCTTGAA
GATAAAACATGTATT**TAA**ACGCCATCTCATATCATGGACTCCGAAGTAGCCTGTTGCCTCC
AAATTTGCCACTTGAATATAATTTTCTTTAAATCGTT

22/75

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60783

><subunit 1 of 1, 330 aa, 1 stop

><MW: 36840, pI: 4.84, NX(S/T): 4

MEGAPPGSLALRLLLFVALPASGWLTTGAPEPPPLSGAPQDGIRINVTTLKDDGDISKQQVV
LNITYESGQVYVNDLPVNSGVTRISCQTLIVKNENLENLEEKEYFGIVSVRILVHEWPMTSG
SSLQLIVIQEEVVEIDGKQVQQKDVTEIDILVKNRGVLRRHSNYTLPLEESMLYSISRDSIL
FTLPNLSKKESVSSLQTTTSQYLIRNVETTVDEDVLPGLPETPLRAEPPSSYKVMCQWMEKF
RKDLCRFWSNVFPVFFQFLNIMVVGITGAADVITILKVFFPVSEYKGILQLDKVDVIPVTAI
NLYPDGPEKRAENLEDKTCI

Important features of the protein:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 266-284

Leucine zipper pattern.

amino acids 155-176

N-glycosylation sites.

amino acids 46-49, 64-67, 166-169, 191-194

23/75

FIGURE 24

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62306

<subunit 1 of 1, 263 aa, 1 stop

<MW: 27663, pI: 6.77, NX(S/T): 2

MRPGAPGPLWPLPWGALAWAVGVSSMSGGNPAPGGVCWLQQGQEATCSLVLQTDVTRAEC
ASGNIDTAWSNLTHPGNKINLLGFLGLVHCLPCKDSCDGVCEGPGKACRMLGGRPRCECAPD
CSGLPARLQVCGSDGATYRDECELRAARCRGHPDLSVMYRGRCRKSCHEHVVCPRPQSCVVDQ
TGSAHCVVCRAAPCPVPSSPGQELCGNNNVTYISSCHMRQATCFLGRSIGVRHAGSCAGTPE
EPPGGESAEEEEENFV

Important features:**Signal peptide:**

amino acids 1-20

N-glycosylation sites.

amino acids 73-77, 215-219

Osteonectin domain proteins.

amino acids 97-130, 169-202

FIGURE 25

TGCAGAGCTTGTGGAGGCCATGGGGCGCGTCGTCGCGGAGCTCGTCTCCTCGCTGCTGGGGT
TGTGGCTGTTGCTGTGCAGCTGCGGATGCCCCGAGGGCGCCGAGCTGCGTGCTCCGCCAGAT
AAAATCGCGATTATTGGAGCCGGAATTGGTGGCACTTCAGCAGCCTATTACCTGCGGCAGAA
ATTTGGGAAAGATGTGAAGATAGACCTGTTTGAAAGAGAAGAGGTCGGGGGCCGCCTGGCTA
CCATGATGGTGCAGGGGGCAAGAATACGAGGCAGGAGGTTCTGTTCATCCATCCTTTAAATCTG
CACATGAAACGTTTTGTCAAAGACCTGGGTCTCTCTGCTGTTTCAGGCCTCTGGTGGCCTACT
GGGGATATATAATGGAGAGACTCTGGTATTTGAGGAGAGCAACTGGTTCATAATTAACGTGA
TTAAATTAGTTTGGCGCTATGGATTTCAATCCCTCCGTATGCACATGTGGGTAGAGGACGTG
TTAGACAAGTTCATGAGGATCTACCGCTACCACTCTCATGACTATGCCTTCAGTAGTGTCGA
AAAATTACTTCATGCTCTAGGAGGAGATGACTTCCTTGGAATGCTTAATCGAACACTTCTTG
AAACCTTGCAAAGGCCGGCTTTTCTGAGAAGTTCCTCAATGAAATGATTGCTCCTGTTATG
AGGGTCAATTATGGCCAAAGCACGGACATCAATGCCTTTGTGGGGGCGGTGTCAGTGTCTG
TTCTGATTCTGGCCTTTGGGCAGTAGAAGGTGGCAATAAACTTGTTTGCTCAGGGCTTCTGC
AGGCATCCAAAAGCAATCTTATATCTGGCTCAGTAATGTACATCGAGGAGAAAACAAAGACC
AAGTACACAGGAAATCCAACAAAGATGTATGAAGTGGTCTACCAAATTGGAAGTGAAGTCTG
TTCAGACTTCTATGACATCGTCTTGGTGGCCACTCCGTTGAATCGAAAAATGTCGAATATTA
CTTTTCTCAACTTTGATCCTCCAATTGAGGAATTCATCAATATTATCAACATATAGTGACA
ACTTTAGTTAAGGGGGAATTGAATACATCTATCTTTAGCTCTAGACCCATAGATAAATTTGG
CCTTAATACAGTTTTTAACCACTGATAATTGAGATTTGTTTCATTAACAGTATTGGGATTGTGC
CCTCTGTGAGAGAAAAGGAAGATCCTGAGCCATCAACAGATGGAACATATGTTTGGAAGATC
TTTTCCCAAGAACTCTTACTAAAGCACAAATTTTAAAGCTCTTCTGTCTCTATGATTATGC
TGTGAAGAAGCCATGGCTTGCAATATCCTCACTATAAGCCCCCGGAGAAATGCCCTCTATCA
TTCTCCATGATCGACTTTATTACCTCAATGGCATAGAGTGTGCAGCAAGTGCCATGGAGATG
AGTGCCATTGCAGCCACACGCTGCACTCCTTGCCATACACGCTGGAACGGGCACACAGA
CATGATTGATCAGGATGGCTTATATGAGAACTTAAACTGAACTATTGAAGTGACACACTCC
TTTTTCCCCTCCTAGTTCCAAATGACTATCAGTGGCAAAAAAGAACAAATCTGAGCAGAGA
TGATTTTGAACCAGATATTTTGCCATTATCATTGTTTAAATAAAAGTAATCCCTGCTGGTCAT
AGGAAAAAAAAAAAAA

FIGURE 27

CATTTCCAACAAGAGCACTGGCCAAGTCAGCTTCTTCTGAGAGAGTCTCTAGAAGAC**ATG**AT
GCTACACTCAGCTTTGGGTCTCTGCCTCTTACTCGTCACAGTTTCTTCCAACCTTGCCATTG
CAATAAAAAAGGAAAAGAGGCCTCCTCAGACACTCTCAAGAGGATGGGGAGATGACATCACT
TGGGTACAACTTATGAAGAAGGTCTCTTTTATGCTCAAAAAAGTAAGAAGCCATTAATGGT
TATTCATCACCTGGAGGATTGTCAATACTCTCAAGCACTAAAGAAAGTATTTGCCCAAATG
AAGAAATACAAGAAATGGCTCAGAATAAGTTCATCATGCTAAACCTTATGCATGAAACCACT
GATAAGAATTTATCACCTGATGGGCAATATGTGCCTAGAATCATGTTTGTAGACCCTTCTTT
AACAGTTAGAGCTGACATAGCTGGAAGATACTCTAACAGATTGTACACATATGAGCCTCGGG
ATTTACCCCTATTGATAGAAAACATGAAGAAAGCATTAAGACTTATTCAGTCAGAGCTA**TAA**
GAGATGATGGAAAAAAGCCTTCACTTCAAAGAAGTCAAATTTTCATGAAGAAAACCTCTGGCA
CATTGACAAATACTAAATGTGCAAGTATATAGATTTTGTAAATATTACTATTTAGTTTTTTTA
ATGTGTTTGCAATAGTCTTATTAAAATAAATGTTTTTTTAAATCTGA

28/75

FIGURE 28

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64896

<subunit 1 of 1, 166 aa, 1 stop

<MW: 19171, pI: 8.26, NX(S/T): 1

MMLHSALGLCLLLVTVSSNLAIKKEKRPPQTLSRGWGDDITWVQTYEEGLFYAQSKKPL
MVIHHLEDCQYSQALKKVFAQNEEIQEMAQNKFIMLNLMHETTDKNLSPDGQYVPRIMFVDP
SLTVRADIAGRYSNRLYTYEPRDLPLLIENMKKALRLIQSEL

Important features:

Signal peptide:

amino acids 1-23

N-myristoylation site.

amino acids 51-57

FIGURE 29

TAAACAGCTACAATATTCCAGGGCCAGTCACTTGCCATTTCTCATAACAGCGTCAGAGAGA
AAGAACTGACTGAAACGTTTGAG**ATGA**AAGAAAGTTCTCCTCCTGATCACAGCCATCTTGCCA
GTGGCTGTTGGTTTCCCAGTCTCTCAAGACCAGGAACGAGAAAAAAGAAGTATCAGTGACAG
CGATGAATTAGCTTCAGGGTTTTTTGTGTTCCCTTACCCATATCCATTTGCCCCACTTCCAC
CAATTCCATTTCCAAGATTTCCATGGTTTAGACGTAATTTTCCTATTCCAATACCTGAATCT
GCCCCTACAACCTCCCCTTCCTAGCGAAAAG**TAA**ACAAGAAGGATAAGTCACGATAAACCTGG
TCACCTGAAATTGAAATTGAGCCACTTCCTTGAAGAATCAAAATTCCTGTTAATAAAAGAAA
AACAAATGTAATTGAAATAGCACACAGCATTCTCTAGTCAATATCTTTAGTGATCTTCTTTA
ATAAACATGAAAGCAAAGATTTTGGTTTCTTAATTTCCACA

30/75

FIGURE 30

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71290
><subunit 1 of 1, 85 aa, 1 stop
><MW: 9700, pI: 9.55, NX(S/T): 0
MKKVLLLLITAILAVAVGFPVSQDQEREKRSISDSDELASGFFVFPYPYPFRPLPPIPFPRFP
WFRRNFPIPIPIESAPTTPLPSEK
```

Important features of the protein:

Signal peptide:

amino acids 1-17

Homologous region to B3-hordein:

amino acids 47-85

FIGURE 31

CGGACGCGTGGGCGGGCGCGCCGGGAGGGACCGGCGGCGGC**ATG**GGGCCGGGGGCCCTGGGAT
GCGGGCCCCGTCTCGCCGCCTGCTGCCGCTGTTGCTGCTGCTCGGCCTGGCCCGCGGCGCCGC
GGGAGCGCCGGGCCCCGACGGTTTAGACGTCTGTGCCACTTGCCATGAACATGCCACATGCC
AGCAAAGAGAAGGGAAGAAGATCTGTATTTGCAACTATGGATTTGTAGGGAACGGGAGGACT
CAGTGTGTTGATAAAAATGAGTGCCAGTTTGGAGCCACTCTTGTCTGTGGGAACACACATC
TTGCCACAACACCCCCGGGGCTTCTATTGCATTTGCCTGGAAGGATATCGAGCCACAAACA
ACAACAAGACATTCAATCCCAACGATGGCACCTTTTGTACAGACATAGATGAGTGTGAAGTT
TCTGGCCTGTGCAGGCATGGAGGGCGATGCGTGAACACTCATGGGAGCTTTGAATGCTACTG
TATGGATGGATACTTGCCAAGGAATGGACCTGAACCTTTCCACCCGACCCAGCATGCCACAT
CATGCACAGAAATAGACTGTGGTACCCCTCCTGAGGTTCCAGATGGCTATATCATAGAAAT
TATACGTCTAGTCTGGGCAGCCAGGTTTCGTTATGCTTGCAGAGAAGGATTCTTCAGTGTTC
AGAAGATACAGTTTCAAGCTGCACAGGCCTGGGCACATGGGAGTCCCCAAAATTACATTGCC
AAGAGATCAACTGTGGCAACCCCTCCAGAAATGCGGCACGCCATCTTGGTAGGAAATCACAGC
TCCAGGCTGGGCGGTGTGGCTCGCTATGTCTGTCAAGAGGGCTTTGAGAGCCCTGGAGGAAA
GATCACTTCTGTTTGCACAGAGAAAGGCACCTGGAGAGAAAGTACTTTAACATGCACAGAAA
TTCTGACAAAGATTAATGATGTATCACTGTTTAAATGATACCTGTGTGAGATGGCAAATAAAC
TCAAGAAGAATAAACCCCAAGATCTCATATGTGATATCCATAAAAGGACAACGGTTGGACCC
TATGGAATCAGTTCGTGAGGAGACAGTCAACTTGACCACAGACAGCAGGACCCCAAGAGTGT
GCCTAGCCCTGTACCCAGGCACCAACTACACCGTGAACATCTCCACAGCACCTCCCAGGCGC
TCGATGCCAGCCGTCATCGGTTTCCAGACAGCTGAAGTTGATCTCTTAGAAGATGATGGAAG
TTTCAATATTTCAATATTTAATGAAACTTGTGTTGAAATTGAACAGGCGTTCTAGGAAAGTTG
GATCAGAACACATGTACCAATTTACCGTTCTGGGTGAGAGGTGGTATCTGGCTAACTTTTCT
CATGCAACATCGTTTAACTTCACAACGAGGGAACAAGTGCCTGTAGTGTGTTTGGATCTGTA
CCCTACGACTGATTATACGGTGAATGTGACCCTGCTGAGATCTCCTAAGCGGCACTCAGTGC
AAATAACAATAGCAACTCCCCCAGCAGTAAACAGACCATCAGTAACATTTCAGGATTTAAT
GAAACCTGCTTGAGATGGAGAAGCATCAAGACAGCTGATATGGAGGAGATGTATTTATTCCA
CATTTGGGGCCAGAGATGGTATCAGAAGGAATTTGCCCAGGAAATGACCTTTAATATCAGTA
GCAGCAGCCGAGATCCCGAGGTGTGCTTGACCTACGTCCGGGTACCAACTACAATGTCAGT
CTCCGGGCTCTGTCTTCGGAACCTCCTGTGGTCACTCTCCCTGACAACCCAGATAACAGAGCC
TCCCCCTCCCGGAAGTAGAATTTTTTACGGTGCACAGAGGACCTCTACCACGCCCTCAGACTGA
GGAAAGCCAAGGAGAAAAATGGACCAATCAGTTTCATATCAGGTGTTAGTGCTTCCCCCTGGCC
CTCCAAAGCACATTTTCTTGTGATTCTGAAGGCGCTTCCTCCTTCTTTAGCAACGCCCTCTGA
TGCTGATGGATACGTGGCTGCAGAACTACTGGCCAAAGATGTTCCAGATGATGCCATGGAGA
TACCTATAGGAGACAGGCTGTACTATGGGGAATATTATAATGCACCCCTTGAAAAGAGGGAGT
GATTACTGCATTATATTACGAATCACAAGTGAATGGAATAAGGTGAGAAGACACTCCTGTGC
AGTTTGGGCTCAGGTGAAAGATTTCGTCACTCATGCTGCTGCAGATGGCGGGTGTGGACTGG
GTTCCCTGGCTGTTGTGATCATTCTCACATTCCTCTCCTTCTCAGCGGTG**TGAT**GGCAGATG
GACACTGAGTGGGGAGGATGCACTGCTGCTGGGCAGGTGTTCTGGCAGCTTCTCAGGTGCC
GCACAGAGGCTCCGTGTGACTTCCGTCCAGGGAGCATGTGGGCCTGCAACTTTCTCCATTCC
CAGCTGGGCCCCATTCTGGATTTAAGATGGTGGCTATCCCTGAGGAGTCACCATAAGGAGA
AACTCAGGAATTCTGAGTCTTCCCTGCTACAGGACCAGTTCTGTGCAATGAACTTGAGACT
CCTGATGTACACTGTGATATTGACCGAAGGCTACATACAGATCTGTGAATCTTGGCTGGGAC
TTCTCTGAGTGTGCTGAGGGTCACTCCTCTAGACATTGACTGCAAGAGAATCTCTGCA
ACCTCCTATATAAAAGCATTTCTGTTAATTCATTGAGAATCCATTCTTTACAATATGCAGTG
AGATGGGCTTAAGTTTGGGCTAGAGTTTGACTTTATGAAGGAGGTGATTGAAAAAGAGAAC
GTGACGTAGGCAAATGTTTCAAGCACTTTAGAAACAGTACTTTTCCCTATAATTAGTTGATAT
ACTAATGAGAAAATATACTAGCCTGGCCATGCCAATAAGTTTCCCTGCTGTGTCTGTTAGGCA
GCATTGCTTTGATGCAATTTCTATTGTCCTATATATTCAAAGTAATGTCTACATTCCAGTA
AAAATATCCCGTAATTAAAAA

FIGURE 33

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGGATG
GGCTGTCTCTGGGGTCTGGCTCTGCCCCTTTTCTTCTTCTGCTGGGAGGTTGGGGTCTCTGG
GAGCTCTGCAGGCCCCAGCACCCGCGAGAGCAGACACTGCGATGACAACGGACGACACAGAAG
TGCCCGCTATGACTCTAGCACCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG
ACCTCTTCTAGGGCCTCAACCCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA
GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTACAAAAACATCTCCCAACTTCATGGTG
TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC
ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA
TGACAGCTCTGAAGAGGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA
CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCTCTTCCGACGGCCCCCATCCAGTCATC
ACCCCGTACGGGCCTCAGAGAGCAGCGCTCTTCCGACGGCCCCCATCCAGTCATCACCCC
GTCACGGGCCTCAGAGAGCAGCGCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT
GGTCCCCGGGATCTGATGTCACTCTCCTCGCTGAAGCCCTGGTGACTGTACAAACATCGAG
GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA
CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC
CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC
ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA
CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA
CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC
AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTGAGCAGTGGGCAAAACAACCTTC
CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC
CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGGCCCTTCCCCACCAGCAGGGACCCCT
CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA
GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCCACGACTGCCCGGAC
GAGGCCGACCACAGACGTGAGTGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG
TGGCTTCCCCGGAAGACCTCACTGACCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC
CGGGAACTCCACGCCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA
ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCTTAGCCTG
GGCCCCACCGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCTATG
AAGGGCAGCATGTCCAAGCCCCTAACCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC
CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCAGAGGTGTCCTTGGACTCACCTTGG
CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA
TTAAAATTCACTCAGTGTGGCCCAAAAAA

FIGURE 34

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108722

><subunit 1 of 1, 482 aa, 1 stop

><MW: 49060, pI: 4.74, NX(S/T): 4

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTL
ETSSRASTPAGPIPEAETRGAKRISPAETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
TTVQTITGSDPEEAI FDTLCTDDSSEEAKTLTMDILTLAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTLGALVTVSRNPLEETSALSVETPSY
VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPTSRD
PLPSVPPTTTNSSRGTNSTLAKITTSAKTTMKPQQPRPRLPGRGRPQT

Important features of the protein:**Signal peptide:**

amino acids 1-25

N-glycosylation sites.

amino acids 252-256, 445-449, 451-455

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-88

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284,
314-320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 35

GCCTCTGAATTGTTGGGCAGTCTGGCAGTGGAGCTCTCCCCGGTCTGACAGCCACTCCAGAG
GCCATGCTTCGTTTCTTGCCAGATTTGGCTTTCAGCTTCCTGTTAATTCTGGCTTTGGGCCA
GGCAGTCCAATTTCAAGAATATGTCTTTCTCCAATTTCTGGGCTTAGATAAGGCGCCTTCAC
CCCAGAAGTTCACACCTGTGCCTTATATCTTGAAGAAAATTTTCCAGGATCGCGAGGCAGCA
GCGACCACTGGGGTCTCCCGAGACTTATGCTACGTAAAGGAGCTGGGCGTCCGCGGGAATGT
ACTTCGCTTTCTCCCAGACCAAGGTTTCTTTCTTTACCCAAAGAAAATTTCCCAAGCTTCCT
CCTGCCTGCAGAAGCTCCTCTACTTTAACCTGTCTGCCATCAAAGAAAGGGAACAGTTGACA
TTGGCCCAGCTGGGCCTGGACTTGGGGCCCAATTCTTACTATAACCTGGGACCAGAGCTGGA
ACTGGCTCTGTTCTGGTTCAGGAGCCTCATGTGTGGGGCCAGACCACCCCTAAGCCAGGTA
AAATGTTTGTGTTGCGGTCAGTCCCATGGCCACAAGGTGCTGTTCACTTCAACCTGCTGGAT
GTAGCTAAGGATTGGAATGACAACCCCCGGAAAAATTTTCGGGTATTCTGAGATACTGGT
CAAAGAAGATAGAGACTCAGGGGTGAATTTTCAGCCTGAAGACACCTGTGCCAGACTAAGAT
GCTCCCTTCATGCTTCCCTGCTGGTGGTGACTCTCAACCCTGATCAGTGCCACCCTTCTCGG
AAAAGGAGAGCAGCCATCCCTGTCCCCAAGCTTTCTTGTAAGAACCTCTGCCACCGTCACCA
GCTATTCATTAACCTCCGGGACCTGGGTTGGCACAAGTGATCATTGCCCCCAAGGGGTTCA
TGGCAAATTACTGCCATGGAGAGTGTCCCTTCTCACTGACCATCTCTCTCAACAGCTCCAAT
TATGCTTTCATGCAAGCCCTGATGCATGCCGTTGACCCAGAGATCCCCCAGGCTGTGTGTAT
CCCCACCAAGCTGTCTCCCATTTCCATGCTCTACCAGGACAATAATGACAATGTCATTCTAC
GACATTATGAAGACATGGTAGTCGATGAATGTGGGTGTGGGTAGGATGTCAGAAATGGGAAT
AGAAGGAGTGTTCTTAGGGTAAATCTTTTAATAAACTACCTATCTGGTTTATGACCACTTA
GATCGAAATGTC

FIGURE 36

MLRFLPDLAFFSFLILALGQAVQFQEYVFLQFLGLDKAPSPQKFQVPYILKKIFQDREAAA
TTGVSRDLCYVKELGVRGNVLRFLPDQGFFLYPKKISQASSCLQKLLYFNLSAIKEREQLTL
AQLGLDLGPNSYYNLGPELELALFLVQEPHVWGQTTPKPGKMFVLRVWPWPQGAVHFNLLDV
AKDWNDNPRKNFGLFLEILVKEDRDSGVNFQPEDTCARLRCSLHASLLVVTLNPDQCHPSRK
RRAAIPVPKLSCKNLCHRHQLFINFRDLGWHKWIIAPKGFMANYPCHGECPPFSLTISLNSSNY
AFMQALMHAVDPEIPQAVCIPTKLSPISMLYQDNNDNVILRHYEDMVVDECGCG

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 112-116, 306-310

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 96-100

N-myristoylation site.

amino acids 77-83

TGF-beta family proteins.

amino acids 264-299, 327-341, 345-364

37/75

FIGURE 37

CACTTTCTCCCTCTCTTCCTTTACTTTTCGAGAAACCGCGCTTCCGCTTCTGGTCGCAGAGAC
CTCGGAGACCGCGCCGGGGAGACGGAGGTGCTGTGGGTGGGGGGACCTGTGGCTGCTCGTA
CCGCCCCCACCCTCCTCTTCTGCACTGCCGTCTCCGGAAGACCTTTTCCCCTGCTCTGTT
TCCTTCACCGAGTCTGTGCATCGCCCCGGACCTGGCCGGGAGGAGGCTTGGCCGGCGGGAGA
TGCTCTAGGGGGCGGCGCGGGAGGAGCGGCCGGCGGGACGGAGGGCCCCGGCAGGAAG**ATG**GGC
TCCCGTGGACAGGGACTCTTGCTGGCGTACTGCCTGCTCCTTGCCCTTTCCTCTGGCCTGGT
CCTGAGTCGTGTGCCCCATGTCCAGGGGGAACAGCAGGAGTGGGAGGGGACTGAGGAGCTGC
CGTCGCCTCCGGACCATGCCGAGAGGGCTGAAGAACAACATGAAAAATACAGGCCCAGTCAG
GACCAGGGGGCTCCCTGCTTCCCGGTGCTTGCGCTGCTGTGACCCCGGTACCTCCATGTACCC
GGCGACCGCCGTGCCCCAGATCAACATCACTATCTTGAAAGGGGAGAAGGGTGACCGCGGAG
ATCGAGGCCCTCAAAGGGAAATATGGCAAACAGGCTCAGCAGGGGGCCAGGGGCCACACTGGA
CCCAAAGGGCAGAAGGGCTCCATGGGGGGCCCTGGGGAGCGGTGCAAGAGCCACTACGCCGC
CTTTTCGGTGGGCGGGAAGAAGCCCATGCACAGCAACCACTACTACCAGACGGTGATCTTCG
ACACGGAGTTCGTGAACCTCTACGACCACTTCAACATGTTACCGGCAAGTTCTACTGCTAC
GTGCCCCGGCCTCTACTTCTTCAGCCTCAACGTGCACACCTGGAACCAGAAGGAGACCTACCT
GCACATCATGAAGAACGAGGAGGAGGTGGTGATCTTGTTTCGCGCAGGTGGGCGACCGCAGCA
TCATGCAAAGCCAGAGCCTGATGCTGGAGCTGCGAGAGCAGGACCAGGTGTGGGTACGCCTC
TACAAGGGCGAACGTGAGAACGCCATCTTCAGCGAGGAGCTGGACACCTACATCACCTTCAG
TGGCTACCTGGTCAAGCACGCCACCGAGCCC**TAG**CTGGCCGGCCACCTCCTTTCTCTCGCC
ACCTTCCACCCCTGCGCTGTGCTGACCCACCGCCTCTTCCCGATCCCTGGACTCCGACTC
CCTGGCTTTTGGCATTTCAGTGAGACGCCCTGCACACACAGAAAGCCAAAGCGATCGGTGCTCC
CAGATCCCGCAGCCTCTGGAGAGAGCTGACGGCAGATGAAATCACCAGGGCGGGGCACCCGC
GAGAACCCTCTGGGACCTTCCGCGGCCCTCTCTGCACACATCCTCAAGTGACCCCGCACGGC
GAGACGCGGGTGGCGGCAGGGCGTCCCAGGGTGCGGCACCGCGGCTCCAGTCTTTGGAATA
ATTAGGCAAATTCTAAAGGTCTCAAAGGAGCAAAGTAAACCGTGGAGGACAAAGAAAAGGG
TTGTTATTTTTGTCTTTCCAGCCAGCCTGCTGGCTCCCAAGAGAGAGGCCTTTTCAGTTGAG
ACTCTGCTTAAGAGAAGATCCAAAGTTAAAGCTCTGGGGTCAGGGGAGGGGGCCGGGGGCAGG
AAACTACCTCTGGCTTAATTCTTTTAAAGCCACGTAGGAACCTTTCTTGAGGGATAGGTGGACC
CTGACATCCCTGTGGCCTTGCCCAAGGGCTCTGCTGGTCTTTCTGAGTCACAGCTGCGAGGT
GATGGGGGGCTGGGGCCCCAGGCGTCAGCCTCCCAGAGGGACAGCTGAGCCCCCTGCCTTGGC
TCCAGGTTGGTAGAAGCAGCCGAAGGGCTCCTGACAGTGGCCAGGGACCCCTGGGTCCCCCA
GGCCTGCAGATGTTTCTATGAGGGGCAGAGCTCCTTGGTACATCCATGTGTGGCTCTGCTCC
ACCCCTGTGCCACCCCAAGAGCCCTGGGGGGTGGTCTCCATGCCTGCCACCCTGGCATCGGCT
TTCTGTGCCGCCTCCACACAAATCAGCCCCAGAAGGCCCGGGGCCTTGGCTTCTGTTTTT
TATAAAACACCTCAAGCAGCACTGCAGTCTCCCATCTCCTCGTGGGCTAAGCATCACCGCTT
CCACGTGTGTTGTGTTGGTTGGCAGCAAGGCTGATCCAGACCCCTTCTGCCCCACTGCCCT
CATCCAGGCCTCTGACCAGTAGCCTGAGAGGGGCTTTTTCTAGGCTTCAGAGCAGGGGAGAG
CTGGAAGGGGCTAGAAAGCTCCCGCTTGTCTGTTTCTCAGGCTCCTGTGAGCCTCAGTCCTG
AGACCAGAGTCAAGAGGAAGTACACGTCCCAATCACCCGTGTCAGGATTCACCTCTCAGGAGC
TGGGTGGCAGGAGAGGCAATAGCCCTGTGGCAATTGCAGGACCAGCTGGAGCAGGGTTGCG
GTGTCTCCACGGTGCTCTCGCCCTGCCCATGGCCACCCAGACTCTGATCTCCAGGAACCC
ATAGCCCTCTCCACCTCACCCATGTTGATGCCCAGGGTCACTCTTGCTACCCGCTGGGCC
CCCAAACCCCGCTGCCTCTCTTCCTTCCCCCATCCCCACCTGGTTTTGACTAATCCTGC
TTCCTCTCTGGGCCTGGCTGCCGGGATCTGGGGTCCCTAAGTCCCTCTCTTAAAGAACTT
CTGCGGGTCAGACTCTGAAGCCGAGTTGCTGTGGGCGTGCCCGGAAGCAGAGCGCCACACTC
GCTGCTTAAGCTCCCCCAGCTCTTTCCAGAAAACATTAAACTCAGAATTGTGTTTTCAA

FIGURE 38

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41234
><subunit 1 of 1, 281 aa, 1 stop
><MW: 31743, pI: 6.83, NX(S/T): 1
MGSRGQGLLLAYCLLLAFASGLVLSRVPHVQGEQQEWEGTEELPSPPDHAERAEEQHEKYRP
SQDQGLPASRCLRCCDPGTSMPATAVPQINITILKGEKGDRGDRGLQGKYGKTGSAGARGH
TGPKGQKGSMGAPGERCKSHYAAFSVGRKKPMHSNHYYQTVIFDTEFVNLYDHFNMFTGKFY
CYVPGLYFFSLNVHTWNQKETYLHIMKNEEEVVILFAQVGDERSIMQSQSLMLELREQDQVWV
RLYKGERENAIIFSEELDTYITFSGYLVKHATEP
```

Signal sequence.

amino acids 1-25

N-glycosylation site.

amino acids 93-97

N-myristoylation sites.

amino acids 7-13, 21-27, 67-73, 117-123, 129-135

Amidation site.

amino acids 150-154

Cell attachment sequence.

amino acids 104-107

FIGURE 39

GAATTCGGCACGAGGGAAGAAGAGAAAGAAAATCTCCGGGGCTGCTGGGAGCATATAAAGAA
GCCCTGTGGCCTTGCTGGTTTTTACCATCCAGACCAGAGTCAGGCCACAGACGGAC**ATGG**CTG
CTCAAGGCTGGTCCATGCTCCTGCTGGCTGTCCTTAACCTAGGCATCTTCGTCCGTCCCTGT
GACACTCAAGAGCTACGATGTCTGTGTATTTCAGGAACACTCTGAATTCATTCCTCTCAAAC
CATTAAAAATATAATGGTGATATTCGAGACCATTTACTGCAACAGAAAGGAAGTGATAGCAG
TCCCCAAAAATGGGAGTATGATTTGTTTGGATCCTGATGCTCCATGGGTGAAGGCTACTGTT
GGCCCAATTACTAACAGGTTCCCTACCTGAGGACCTCAAACAAAAGGAATTTCCACCGGCAAT
GAAGCTTCTGTATAGTGTTGAGCATGAAAAGCCTCTATATCTTTCATTTGGGAGACCTGAGA
ACAAGAGAATATTTCCCTTTCCAATTCGGGAGACCTCTAGACACTTTGCTGATTTAGCTCAC
AACAGTGATAGGAATTTTCTACGGGACTCCAGTGAAGTCAGCTTGACAGGCAGTGATGCC**TA**
AAAGCCACTCATGAGGCAAAGAGTTTCAAGGAAGCTCTCCTCCTGGAGTTTGGCGTTCTCA
TTCTTATACTCTATTTCCCGCGTTAGTCTGGTGTATGGATCTATGAGCTCTCTTTTAATATTT
TATTATAAATGTTTTATTTACTTAACTTCCTAGTGAATGTTACAGGTGACTGCTCCCCCAT
CCCCATTTCTTGATATTACATATAATGGCATCATATACCCCTTTATTGACTGACAACTACT
CAGATTGCTTAACATTTTGTGCTTCAAAGTCTTATCCCACTCCACTATGGGCTGTTACAGAG
TGCATCTCGGTGTAGAGCAAGGCTCCTTGTCTTCAGTGCCCCAGGGTGAAATACTTCTTTGA
AAAATTTTCATTCATCAGAAAATCTGAAATAAAAAATATGTCTTAATTGAG

40/75

FIGURE 40

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73838
><subunit 1 of 1, 167 aa, 1 stop
><MW: 19091, pI: 7.48, NX(S/T): 1
MAAQGWSMLLLAVLNLGIFVRPCDTQELRCLCIQEHSEFIPLKLIK NIMVIFETIYCNRKEV
IAVPKNGSMICLDPDAPWVKATVGPITNRFLPEDLKQKEFPPAMKLLYSVEHEKPLYLSFGR
PENKRIFPFPIRETSRHFADLAHNSDRNFLRDSSEVSLTGSDA
```

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 68-72

N-myristoylation site.

amino acids 69-75

Small cytokines (intercrine/chemokine) C-x-C subfamily signature

amino acids 40-85

FIGURE 41

CAGAC**ATG**GCTCAGTCACTGGCTCTGAGCCTCCTTATCCTGGTTCTGGCCTTTGGCATCCCC
AGGACCCAAGGCAGTGATGGAGGGGCTCAGGACTGTTGCCTCAAGTACAGCCAAAGGAAGAT
TCCCGCCAAGGTTGTCCGCAGCTACCGGAAGCAGGAACCAAGCTTAGGCTGCTCCATCCCAG
CTATCCTGTTCTTGCCCCGCAAGCGCTCTCAGGCAGAGCTATGTGCAGACCCAAAGGAGCTC
TGGGTGCAGCAGCTGATGCAGCATCTGGACAAGACACCATCCCCACAGAAACCAGCCCAGGG
CTGCAGGAAGGACAGGGGGGCCTCCAAGACTGGCAAGAAAGGAAAGGGCTCCAAAGGCTGCA
AGAGGACTGAGCGGTCACAGACCCCTAAAGGGCCA**TAG**CCCAGTGAGCAGCCTGGAGCCCTG
GAGACCCACACAGCCTCACCAGCGCTTGAAGCCTGAACCCAAGATGCAAGAAGGAGGCTATG
CTCAGGGGGCCCTGGAGCAGCCACCCCATGCTGGCCTTGCCACACTCTTTCTCCTGCTTTAAC
CACCCCATCTGCATTCCCAGCTCTACCCTGCATGGCTGAGCTGCCCACAGCAGGCCAGGTCC
AGAGAGACCGAGGAGGGAGAGTCTCCAGGGAGCATGAGAGGAGGCAGCAGGACTGTCCCCT
TGAAGGAGAATCATCAGGACCCTGGACCTGATACGGCTCCCCAGTACACCCACCTCTTCCT
TGTAATATGATTTATACCTAACTGAATAAAAAGCTGTTCTGTCTTCCCNCCCCA

FIGURE 42

><MW: 14646, pI: 10.45, NX(S/T): 0

MAQSLALSLILVLAFGIPTQGS DGG AQDCCLKYSQRKIPAKVVRSYRKQEPSLGCSIPAI
LFLPRKRSQAELCADPKELWVQQLMQHLDKTPSPQKPAQGCRKDRGASKTGKKKGKSGKCKR
TERSQTPKGP

Important features of the protein:

Signal peptide:

amino acids 1-17

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 67-71

N-myristoylation sites.

amino acids 17-23, 23-29, 27-33, 108-114, 118-124, 121-127

Amidation site.

amino acids 112-116

Small cytokines.

amino acids 51-91

43/75

FIGURE 43

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGC**ATGA**AGTTACAGTGTGTTTCCCTTTGGCTC
 CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTT
 CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAAGAGCCATCCAAGCTAAGG
 ACACCTTCCCAAATGTCACCTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
 GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAAGGATCA
 TCAGGAGCCAAACCCCAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA
 TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC
 AATGCCACCAGAGTCATCCATGACAACCTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
 ATCCCTGGGAGAGCTCGACGTCTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
 CAGCT**TGA**TGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT
 GTGGGAGACAGCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGCAGCTGAAAGTCC
 CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGCAAAAAGTCTACTGTGGTAT
 TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
 CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
 GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCATATTTTACCTATGA

FIGURE 44

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282
><subunit 1 of 1, 177 aa, 1 stop
><MW: 20452, pI: 8.00, NX(S/T): 2
MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ
RQCHCRQEATNATRVIHDNYDQLEVHAAIKSLGELDVFLAWINKNHEVMFSA
```

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

FIGURE 45

GCTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG
 CTAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTACCCCGGCTGGGA
 CAAGAAGCCGCCGCTGCCTGCCCGGGCCCGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGG
 GGTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
 TGTCTTGGGCACCTACCCGTGGGGCCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCCGGAG
 CCGCCGCGCCGTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC
 GGCACCTCACAGCCCCGACGCGCATCCCGGTGCGCGCCACGCCTCCCGCACCCCCATCGCCGG
 AGCTGCGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGGTCCACGTATGG
 ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC
 CCACGTGCACTACGGCTGGGGCGACCCCATCCGCTGCGGCACCTGTACACCTCCGGCCCCC
 ACGGGCTCTCCAGCTGCTTCCTGCGCATCCGCTGCGGCAGCGGCGTCTGCGGACCGTGGCCATCAAGGG
 CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG
 CGTGACAGCGTGCGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT
 ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA
 TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA
 CAGAGGCTTTCTTCCACTCTCTCATTTTCTGCCCATGCTGCCCATGGTCCCAGAGGAGCCTG
 AGGACCTCAGGGGCCACTTGGAACTTGACATGTTCTCTTCGCCCCCTGGAGACCGACAGCATG
 GACCCATTTGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCCAGCTTTGAGAAGTAACT
 GAGACCATGCCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG
 TGCTTCTACAAGAACAGTCCCTGAGTCCACGTTCTGTTTACGCTTTAGGAAGAAACATCTAGAA
 GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT
 CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGCCTGGGCCCCCATTCTGCTCCCTCGA
 GGTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC
 TCACTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC
 CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTACAGGAACAGGTGATCCACTCTGTA
 AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
 ACCATTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCACCAG
 GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACCTGAGAATTCCCC
 CTGAGGCCAGTTCTGTCATGGATGCTGTCCTGAGAATAACTTGCTGTCCCGGTGTCACCTGC
 TTCCATCTCCCAGCCCACCAGCCCTCTGCCACCTCACATGCCTCCCCATGGATTGGGGCCT
 CCCAGGCCCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT
 TTGAAGACCCCAAGTCTTGTCAATAACTTGTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC
 CCTTTCCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACA
 TCTCTTATTTTCTTACATTATTTATGCCCCCAAATTATATTTATGTATGTAAGTGAGGTTTG
 TTTTGTATATTAAATGGAGTTTGTGTTGT

FIGURE 46

MRSGCVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWDPIRLRHLYTSGPHGLSSCFLRI
RADGVVDCARGQSAHSLEIKAVALRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFE
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD
MFSSPLETDSMDPFGLVTGLEAVRSPSF EK

Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

FIGURE 47

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCG**ATG**GGGACAAA
GGCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTTACAGTGCACCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAACCTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT
CTTGGTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTTCCTGGTG**TGA**GCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT
CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC
TACACCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC
ATCCTCCTTCATGCCCTCCCTCCCTTTCTTACCCTGCTGAGTGGCCTGGAACCTGTTTTAAA
GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAAATGGCGGGGGTTCGCAGGAATCTGCACTCAACTGCCACCTGGC
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCTTGTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG
TGATGACACTGGGGTCCCTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACCTGGAGGCTGGGCGCAGTGGCTCACGCCTG
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTGCGGAGTTCGGGATCAGCCTGACCA
ACATGGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

FIGURE 48

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSLAYSGFSSPRVEW
 KFDQGDTTTRLVCYNNKITASYEDRVTFLLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV
 KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPEYTWFKDGIVMPTNPKSTRAFSNS
 SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVA AVLVTLI
 LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
 193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

[illegible]

FIGURE 50

MGAARLLPNLTLCLQLLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRI
SATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLE
NNYTAFQONARHEGWFMATRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEF
VGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 51

52/75

FIGURE 52

MKTIQPKMHNSISWAI FTGLAALCLFQGV PVRSGDATFPKAMDNVTVRQGESATL RCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVLLSNTQTQYSIEIQNV DVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLT CIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYEC SASNDVAAPV VRRVKVTVNYPPYISEAKGTGVPVGQKGT LQCEASAV
PSAEFQWYKDDKRLIEGKKG VKVENRPFLSKLI FFNVSEHDYGN YTCVASNKLGH TNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

[illegible]

54/75

FIGURE 54

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212

><subunit 1 of 1, 440 aa, 1 stop

><MW: 42208, pI: 6.36, NX(S/T): 1

MKFQGPLACLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
AGSKVSEALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
ADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGGLGGQGNPGGLGTPWVHGYPGNSAGSFGM
NPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGS
GSQSGSSGSGSNGDNNNGSSSGSSSGSSSGSSSGSSSGSSSGSSSGSSSGSSSGNSGSGSRGDSGSESSW
GSSTGSSSGNHGGSGGGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRLL
GGSGDNRYRGQSSWGSGGGDAVGGVNTVNSETPGMFNFDTFWKNFKSKLGFINWDINKDQ
RSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 55

AGCCAGGCAGCACATCACAGCGGGAGGAGCTGTCCCAGGTGGCCCAGCTCAGCA**ATGG**CAAT
GGGGGTCCCCAGAGTCATTCTGCTCTGCCTCTTTGGGGCTGCGCTCTGCCTGACAGGGTCCC
AAGCCCTGCAGTGCTACAGCTTTGAGCACACCTACTTTGGCCCCTTTGACCTCAGGGCCATG
AAGCTGCCCAGCATCTCCTGTCCTCATGAGTGCTTTGAGGGCTATCCTGTCTCTGGACACCGG
GTATCGCGCGCCGGTGACCCTGGTGCGGAAGGGCTGCTGGACCGGGCCTCCTGCGGGGCCAGA
CGCAATCGAACCCGGACGCGCTGCCGCCAGACTACTCGGTGGTGCGCGGCTGCACAACCTGAC
AAATGCAACGCCCACCTCATGACTCATGACGCCCTCCCCAACCTGAGCCAAGCACCCGACCC
GCCGACGCTCAGCGGCGCCGAGTGCTACGCCTGTATCGGGGTCCACCAGGATGACTGCGCTA
TCGGCAGGTCCCGACGAGTCCAGTGTCACCAGGACCAGACCGCCTGCTTCCAGGGCAGTGGC
AGAATGACAGTTGGCAATTTCTCAGTCCCTGTGTACATCAGAACCTGCCACCGGCCCTCCTG
CACCACCGAGGGCACCACCAGCCCCCTGGACAGCCATCGACCTCCAGGGCTCCTGCTGTGAGG
GGTACCTCTGCAACAGGAAATCCATGACCCAGCCCTTACCAGTGCTTCAGCCACCACCCCT
CCCCGAGCACTACAGGTCCTGGCCCTGCTCCTCCCAGTCCTCCTGCTGGTGGGGCTCTCAGC
ATAGACCGCCCCCTCCAGGATGCTGGGGACAGGGCTCACACACCTCATTCCTTGCTGCTTCAGC
CCCTATCACATAGCTCACTGGAAAATGATGTTAAAGTAAGAATTGCAAAA

FIGURE 56

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA86576
><subunit 1 of 1, 251 aa, 1 stop
><MW: 26935, pI: 7.42, NX(S/T): 2
MAMGVPRVILLCLFGAALCLTGSQALQCYSFEHTYFGPFDLRAMKLPSISCPHECFEAILSL
DTGYRAPVTLVRKGCWTGPPAGQTQSNPDALPPDYSVVRGCTTDKCNAHLMTHDALPNLSQA
PDPPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQTACEQGSGRMTVGNFSPVYIIRTCHR
PSCTTEGTTSPWTAIDLQGSCCEGYLCNRKSMTQPFTSASATTPPRALQVLALLLPVLLLVG
LSA
```

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 233-251

N-glycosylation sites.

amino acids 120-124, 174-178

N-myristoylation sites.

amino acids 15-21, 84-90

57/75

FIGURE 57

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCG
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGGCCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGCGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT
TTCCTCTGTACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT
TGACCCCCACGCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT
GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTAT**TGAT**GAC
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCCTATAAGGGTCACCTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCAGGC
CCCCTTGATCTGTACCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAAACTAACATGAAATATGTGTTGTTTTATTGCAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

FIGURE 58

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS
 SQPWEVFPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY
 SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSQCSPRSKPAVQYQ
 WDRQLPSFQTFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA
 VVAGAVVGTTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPWKSSDTISKNGTL
 SSVTSARALRPPHGP RP GALTP T PSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSR
 MGAVPVMVPAQSQAGSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
 262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

59/75

FIGURE 59

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG
GACAGCAGCAAAGAGGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGGAGATTATTTTAC
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT
CATTTTGATTTTGCTGTTTATTTTTTTTTTCTTTTTCTTTTTTCCCACCACATTGTATTTTAT
TTCCGTACTTCAGAAATGAGGCTACAGACCACAAAGTGGCCAGCCATGGGGCTTTTTTCTT
GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC
CTAGTGTGTGCCGCTGCGACAGGAAGTGTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC
TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC
AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC
AATATTCAGACCATTTTACGGGCTGCTCTTGCCAGCTCTTGAAGCTTGAAGAGCTGCACCT
GGATGACAACCTCCATATCCACAGTGGGGGTGGAAGACGGGGCTTCCGGGAGGCTATTAGCC
TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCTGTGGAC
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTTCATATCCGACATGGCCTTCCAGAA
TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG
AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC
CACCTCCTCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT
AAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA
ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC
ACTGCTCGGAATAACCTTGGTTTTGTGACTGCAGTATTAAATGGGTACAGAAATGGCTCAA
ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCTGAACAAGTCCGGG
GGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCTGTCCCACCACGACCCCCGGCCTG
CCTCTCTTACCCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT
TCCAAACCTTAGCAGAAGCTACACGCCTCCAACCTCCTACCACATCGAAACTTCCCACGATTC
CTGACTGGGATGGCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATC
CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTCCACCGTGATGGCATA
CAAACCTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTTCAGGAGCGCATAG
TCAGCGGTGAGAAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT
TGTTTAGTGCCACTGGATGCTTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTTCAGAGGC
CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA
CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT
GTGCTGGTGGTCTTGCTCAGCGTCTTTTGTGCTGGCATATGCACAAAAGGGGCGCTACACCTC
CCAGAAGTGGAATACAACCGGGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA
AGGACAACCTCCATCCTGGAGATGACAGAAACCAGTTTTTCAGATCGTCTCCTTAAATAACGAT
CAACTCCTTAAAGGAGATTTTCACTGTCAGCCCATTTTACACCCCAAATGGGGGCATTAATTA
CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC
ACTGCCATACGTGACAGCCAGAGGGCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT
TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTG
CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTTAAATCTT

FIGURE 60

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP
 EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFFPMNLPKNVRVLHLQENNIQTI
 SRAALAQLLKLEELHLDNNSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELR
 VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD
 LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN
 PWFCDCSIKWVTEWLKYIPSSLNVRGFMCGPEQVRGMARELNMNLLSCPTTTPGLPLFTP
 APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND
 TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL
 DAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLAGLIGGAVIFVLVVL
 LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG
 DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
 522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

FIGURE 61

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAAGTGTATTTACTGCTGCGTTT
TATGTTGGGAATTCCTCTCCTATGGCCTTGTCTTGGAGCAACAGAAAACCTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCGATCTCATTTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTTAGA
CAATGGAAACAATTCTTTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA
TTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA
GTTTGTTCATCAAAGTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATG
AGGCCATTGTACCAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAAGTGGTAATAATGCTCGTCTCCTCTACAGCTTACTTCAAGGCCAGCC
ATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAACAAGTGTATTAATTAACCTTTTTCAGATGTTAATGACAATAAGCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTGAAGAGGAT
GATTCGCAAACATTTGACATTATTACTAATCATGAACTCAAGAAGGAATAGTTATATTTAAA
AAAGAAAGTGGATTTTGGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATC
ATGTTCTCTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCCTTTTATTAAGATCCAG
GTGGAAGATGTTGATGAGCCTCCTCTTTTCTCCTTCCATATTATGTATTTGAAGTTTTTGA
AGAAACCCACAGGGATCATTGTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAAT
CTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTTCAATATCAATGATAATGGTACAATC
ACTACAAGTAAGTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCCTGTATGTGCAAGTCTTAAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAAATGCAGGCTCT
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCACCATT
TTACTTTAATCTATCTGTAGAAAGACACTAACAATTCAAGTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCAATTTGACTAATAGAAGTGGTTTTAACCTTCAAGAAGAACCTGTCTTC
TACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACAAACACCCTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGCTTTCCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTTGCATTATGATCATA
TTTGGGTTTATTTTTTTTGAAGTTTGGGTTTAAAACAACGGAGAAAACAGATTCTATTTCTGA
GAAAAGTGAAGATTTTCAAGAGAGAATATATTCCAATATGATGATGAAGGGGGTGGAGAAGAAG
ATACAGAGGCCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAACGCAAGACT
CGGAAAACCAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA
CAGTGCCATATTCAGGAAATTCATTCTGGAAAAGCTCGAAGAAGCTAATACTGATCCGTGTG
CCCCCTCCTTTTGAATCCCTCCAGACCTACGCTTTTGGAGGGAACAGGGTTCATTAGCTGGATCC
CTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTTAAAAGATTAGCATGCATGTTGGTCTGCAAGTGCAGTCAAATAATTAGG
GCTTTTTTACCATCAAAATTTTTAAAAGTGCTAATGTGTATTGCAACCAATGGTAGTCTTAA
AGAGTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTCTGATTTCC
CTGGAGTAAATACTCCATGGTTATTTTTAAGCTACCTACATGCTGTCATTGAACAGAGATGTG
GGGAGAAATGTAAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAAATAATG
TAGGAAGATATTTAAAGTAGATGAGAGGACACAAGATGTAGTCGATCCTTATGCGATTATAT
CATTATTTACTTAGGAAAGAGTAAAAATACCAAACGAGAAAATTTAAAGGAGCAAAAATTTG
CAAGTCAAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATGAAA
ATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTATTCATGAAGTATTGTTTCCTTTAT
TTAAA

FIGURE 62

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKRGVWNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVS DINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLLQGQPYFSVEPTTG VIRISSKMDRELQDEYVWIIQAKDMIGQPGALSGTTSVLIKLSD
VNDNKPIFKESLYRLTVSESAPTGT SIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE
TQEGIVILKKKVDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYYVFEVFEETPQGSFVGVVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEEHFHYFNLSVEDTNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDGDSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGS LAGSLSSLES AVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSN
```

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518, 516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 63

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCA**ATGGC**
CGCCCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTC
TCTTGGCCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCAGTGCAGGCTTGAC
AAGTCCAACCTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAG
CTTGGCTGATAACAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTA
TGAGTGAGCGCTGCTATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTCC
CCTCAATCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAG
CAACAGGCTAAGCACATGTCATATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAA
AGCTGAAGGACACAGTGAAAAAGCTTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAAGTG
GATTTGCTGTTTATGTCTCTGAGAAATGCCTGCATT**TGACC**AGAGCAAAGCTGAAAAATGAA
TAACTAACCCCTTTCCCTGCTAGAAATAACAATTAGATGCCCCAAAGCGATTTTTTTTAAAC
CAAAGGAAGATGGGAAGCCAACTCCATCATGATGGGTGGATTCCAAATGAACCCCTGCGT
TAGTTACAAAGGAAACCAATGCCACTTTTGTTTATAAGACCAGAAGGTAGACTTTCTAAGCA
TAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTCTATACACAGAAAACAATTTATTT
TTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCATTCCTTTAGGGGAAAAAACCC
CTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAATGTATTTATTATTATTA
TAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTATAGAAACATCATTCG
ATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTATAGAGCTATAAC
ATGTTTATTTGACCTCAATAAACACTTGGATATCCC

64/75

FIGURE 64

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125185
><subunit 1 of 1, 179 aa, 1 stop
><MW: 20011, pI: 8.10, NX(S/T): 3
MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKE
ASLADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLAR
LSNRLSTCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI
```

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

65/75

FIGURE 66

MARFPKADLAAAGVMLLCHFFTDQFQFADGKPGDQILDWQYGVTVQAFPHTEEEVEVD SHAYS
 HRWKRNLDLFLKAVDTNRASVGQDSPEPRSFDTLLLDGQDNNTQIEEDTDHNYIISRIYGPS
 DSASRDLWVNIDQMEKDKVKI HGILSNTHRQAARVNLSFDFPFYGHFLREITVATGGFIYTG
 EVVHRMLTATQYIAPLMANFDPSVSRNSTVRYFDNGTALVVQWDHVHLQDNYNLGSFTFQAT
 LLM DGRIIFGYKEIPVLVTQISSTNHPVKVGLSDAFVVVHRIQQIPNVRRTIYEYHRVELQ
 MSKITNISAVEMTPLPTCLQFNRCGPCVSSQIGFNC SWCSKLQRCSSGFDRHRQDWVDSGCP
 EESKEKMCENTEPVETSSRTTTTVGATTTQFRVLTTTRRAVTSQFPTSLPTEDDTKIALHLK
 DNGASTDDSAAEKKGGLHAGLIIGILILVLIVATAILVTVYMYHHPTSAASIFFIERRPSR
 WPAMKFRRGSGHPAYAEVEPVGEKEGFIVSEQC

Important features of the protein:**Transmembrane domain:**

amino acids 454-478

N-glycosylation sites.

amino acids 103-107, 160-164, 213-217, 221-225, 316-320, 345-349

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 297-301, 492-496, 503-507

N-myristoylation sites.

amino acids 42-48, 100-106, 147-153, 279-285, 397-403, 450-456,
 455-461

FIGURE 67A

GCAGCCCTAGCAGGGATGACATGATGCTGTTGGTGCAGGGTGCCTTGTGCTCGAACCAGTG
GCTGGCGGCGGTGCTCCTCAGCCTGTGCTGCCTGCTACCCCTCCTGCCTCCCGGCTGGACAGA
GTGTGGACTTCCCCTGGGCGGCCGTGGACAACATGATGGTCAGAAAAGGGGACACGGCGGTG
CTTAGGTGTTATTTTGAAGATGGAGCTTCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTAT
TTTTGCGGGAGGTGATAAGTGGTCAGTGGATCCTCGAGTTTCAATTTCAACATTGAATAAAA
GGGACTACAGCCTCCAGATACAGAATGTAGATGTGACAGATGATGGCCCATACACGTGTTCT
GTTCAAGTCAACATACACCCAGAACAATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAA
GATATATGACATCTCAAATGATATGACCGTCAATGAAGGAACCAACGTCCTCTTACTTGTT
TGGCCACTGGGAAACCAGAGCCTTCCATTTCTTGGCGACACATCTCCCCATCAGCAAAACCA
TTTGAAAATGGACAATATTTGGACATTTATGGAATTACAAGGGACCAGGCTGGGGAATATGA
ATGCAGTGCAGGAAAATGATGTGTCAATCCCAGATGTGAGGAAAGTAAAGTTGTTGTCAACT
TTGCTCCTACTATTTCAGGAAATTAAATCTGGCACCGTGACCCCCGGACGCAGTGGCCTGATA
AGATGTGAAGGTGCAGGTGTGCCGCCTCCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCT
CTTCAATGGCCAACAAGGAATTATTATTCAAAATTTTAGCACAAGATCCATTCTCACTGTTA
CCAACGTGACACAGGAGCACTTCGGCAATTATACTTGTGTGGCTGCCAACAAGCTAGGCACA
ACCAATGCGAGCCTGCCTCTTAACCCTCCAAGTACAGCCAGTATGGAATTACCGGGAGCGC
TGATGTTCTTTTCTCCTGCTGGTACCTTGTGTTGACACTGTCCTCTTTCACCAGCATATTCT
ACCTGAAGAATGCCATTCTACAATAAATTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCT
GAAAGTGCTGATGGCTGGATCCAATCTGGTACAGTTTGTTAAAAGCAGCGTGGGATATAATC
AGCAGTGCTTACATGGGGATGATCGCCTTCTGTAGAATTGCTCATTATGTAAATACTTTAAT
TCTACTCTTTTTTTGATTAGCTACATTACCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAG
ACGTGAAAGCTCTGAAATTACTTTTAGAGGATATTAATTGTGATTTTCATGTTTGTAATCTAC
AACTTTTCAAAGCATTTCAGTCATGGTCTGCTAGGTTGCAGGCTGTAGTTTACAAAACGAA
TATTGCAGTGAATATGTGATTCTTTAAGGCTGCAATACAAGCATTTCAGTTCCTGTTCAT
AAGAGTCAATCCACATTTTACAAAGATGCATTTTTTTTTCTTTTTTGATAAAAAAGCAAATAATA
TTGCCTTCAGATTATTTCTTCAAAATATAACACATATCTAGATTTTTCTGCTCGCATGATAT
TCAGGTTTCAGGAATGAGCCTTGTAATATAACTGGCTGTGCAGCTCTGCTTCTCTTCTGT
AAGTTTCAGCATGGGTGTGCCCTTCATACAATAATATTTTTCTCTTTGTCTCCAATAATATAA
AATGTTTTTGCTAAATCTTACAATTTGAAAGTAAAAATAAACCAGAGTGATCAAGTTAAACCA
TACACTATCTCTAAGTAACGAAGGAGCTATTGGACTGTAAAAATCTCTTCTGCACTGACAA
TGGGGTTTGAGAATTTTGCCCCACACTAACTCAGTTCTTGTGATGAGAGACAATTTAATAAC
AGTATAGTAAATATAACCATATGATTTCTTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGA
AATCATTCCTTTTAAATGACAGCACAGTCCACTCAAAGGATTGCCTAGCAATACAGCATCT
TTTCTTTTCACTAGTCCAAGCCAAAAATTTTAAGATGATTTGTGAGAAAGGGCACAAAGTCC
TATCACCTAATATTACAAGAGTTGGTAAGCGCTCATCATTAAATTTTATTTTGTGGCAGCTAA
GTTAGTATGACAGAGGCAGTGCTCCTGTGGACAGGAGCATTTTGCATATTTTCCATCTGAAA
GTATCACTCAGTTGATAGTCTGGAATGCATGTTATATATTTTAAACTTCCAAAATATATTA
TAACAAACATTCTATATCGGTATGTAGCAGACCAATCTCTAAAATAGCTAATTCTTCAATAA
AATCTTTCTATATAGCCATTTTCAGTGCAACAAGTAAATCAAAAAGACCATCCTTTATTT
TTCCTTACATGATATATGTAAGATGCGATCAAATAAAGACAAAACACCAGTGATGAGAATAT
CTTAAGATAAGTAATTATCAAATTATTGTGAATGTTAAATTATTTCTACTATAAAGAAGCAA
AATACATTTTTTGAAGGAAAATGCTGTTACTCTAACATTAATTTACAGGAATAGTTTGATGG
TTTCACTCTTTACTAAAGAAAGGCCATCACCTTGAAAGCCATTTTACAGGTTTGATGAAGTT
ACCAATTTTCAGTACACCTAAATTTCTACAAATAGTCCCCTTTTACAAGTTGTAACAACAAAG
ACCCTATAATAAAATTAGATACAAGAAATTTTGCAGTGGTTATACATATTTGAGATATCTAG
TATGTTGCCCTAGCAGGGATGGCTTAAAAACTGTGATTTTTTTTTCTTCAAGTAAAACCTAGT
CCCAAAGTACATCATAAATCAATTTTAATTAGAAAAATGAATCTTAAATGAGGGGACATAAG
TATACTCTTTCCACAAAATGGCAATAATAAGGCATAAAGCTAGTAAATCTACTAACTGTAAT
AAATGTATGACATTATTTTGATTGATACATTAAAAAAGAGTTTTTTAGAACAAATATGGCATT
TAACTTTATTATTTATTTGCTTTTAAGAAATATTCTTTGTGGAATTGTTGAATAAACTATAA
AATATTATTTTGTATTGCAGCTTTAAAGTGGCACACTCCATAATAATCTACTTACTAGAAAT

FIGURE 67B

AGTGGTGCTACCAACAAAAAATGTTAACCATCAGTACCATTGTTTGGGGAGAAAGAAACAGATC
AAGAATGCATATTATTTCAGTGACCGCTTTCCTAGAGTTAAAATACCTCCTCTTTGTAAGGTT
TGTAGGTAAATTGAGGTATAAACTATGGATGAACCAAATAATTAGTTCAAAGTGTTGTCATG
ATTCCAAATTTGTGGAGTCTGGTGTTTTTACCATAGAATGTGACAGAAGTACAGTCATAGCT
CAGTAGCTATATGTATTTGCCTTTATGTTAGAAGAGACTTTCCTTGAGTGACATTTTTTAAATA
GAGGAGGTATTCACTATGTTTTTCTGTATCACAGCAGCATTCCCTAGTCCTTAGGCCCTCGGA
CAGAGTGAAATCATGAGTATTTATGAGTTCAATATTGTCAAATAAGGCTACAGTATTTGCTT
TTTTGTGTGAATGTATTGCATATAATGTTCAAGTAGATGATTTTACATTTATGGACATATAA
AATGTCTGATTACCCCATTTTATCAGTCCTGACTGTACAAGATTGTTGCAATTTCAGAATAG
CAGTTTTATAAATTGATTTATCTTTTAATCTATAACAATTTGTGTTAGCTGTTCAATTCAGG
ANTATATTTTCTACAAGTTCCTACTTGTGGGACTCCTTTTGTGCCCCCTATTTTTTTTTTAAAG
AAGGAAGAAAGAAAAATAAGTAGCAGTTTAAAAATGAGAATGGAGAGAAAAGAAAAAGAATG
AAAAGGAAAGGCAGTAAAGAGGGGAAAAAAAAGGAAGGATGGAAGGAATGAAGGAAGGAAGGG
AGGAAGGGGAGAAGGTAGGAAGAAAGAAAGGATGAGAGGGAAGGAAGAATCAGAGTATTAGG
GTAGTTAACTTACACATTTGCATTCTTAGTTTTAACTGCAAGTGGTGTAACTATGTTTTTCAA
TGATCGCATTTGAAACATAAGTCCTATTATACCATTAAAGTTCCCTATTATGCAGCAATTATAT
AATAAAAAGTACTGCCCAAGTTATAGTAATGTGGGTGTTTTTGAGACACTAAAAGATTTGAG
AGGGAGAATTTCAAACCTTAAAGCCACTTTTGGGGGGTTTATAACTTAACTGAAAAATTAATG
CTTCATCATAACATTTAAGCTATATCTAGAAAGTAGACTGGAGAACTGAGAAAATTACCCAG
GTAATTCAGGGAAAAAAAATAATATATATATATAAATACCCCTACATTTGAAGTCAGAAA
ACTCTGAAAACTGAATTATCAAAGTCAATCATCTATAATGATCAAATTTACTGAACAATTG
TTAATTTATCCATTGTGCTTAGCTTTGTGACACAGCCAAAAGTTACCTATTTAATCTTTTCA
ATAAAAATTTGTTTTTGAATCCAGAAATGATTTAAAAAGAGGTACAGGTTTTTAACTATTTA
TTGAAGTATGTGGATGTACAGTATTTCAATAGATATGAATATGAATAATGGTATGCCTTAA
GATTCCTTTGAATATGTATTTACTTTTAAAGACTGGAAAAAGCTCTTCCTGTCTTTTAGTAAAA
CATCCATATTTTCATAACCTGATGTAAAAATATGTTGTACTGTTTCCAATAGGTGAATATAAAC
TCAGTTTTATCAATTAA

FIGURE 68

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92259

><subunit 1 of 1, 354 aa, 1 stop

><MW: 38719, pI: 6.12, NX(S/T): 6

MDMMLLVQGACCSNQWLA AVL LSLCCLLP SCLPAGQSVDFPWA AVDNMMVRKGD TAVLR CYL
EDGASKGAWLNRSSII FAGGDKWSV DPRVSISTL NKR DYS LQIQNV DVTDDGPYTCSVQTQH
TPRTMQVHLTVQVPPKIYDISNDMTVNEGTVTLTCLATGKPEPSISWRHISPSAKPFENGQ
YLDIYGITRDQAGEYECSAENDVSFPDVRKVKVVVNFAPTIQEI KSGTVTPGRSGLIRCEGA
GVPPPAFEWYKGEKKLFNGQQGII IQNFSTRSILTVTNVTQEHFGNYTCVAANKLGT TNASL
PLNPPSTAQYGITGSADVLFSCWYLVLTLS SFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 322-343

N-glycosylation sites.

amino acids 73-77, 155-159, 275-279, 286-290, 294-298, 307-311

Tyrosine kinase phosphorylation site.

amino acids 180-188

N-myristoylation sites.

amino acids 9-15, 65-71, 69-75, 153-159, 241-247, 293-299,
304-310, 321-327

Myelin P0 protein.

amino acids 94-123

FIGURE 69

ATAGTAGAAGAATGTCTCTGAAATTACTGGATGAGTTTCAGTCATACTTTCACATGGGCACA
ATTTACATTC AAGCTCCTTATCCTAGGCTAATTTTATATTATGTTAAATCACTTGTTTTTG
TTCTCACGGCTTCCTGCCTGCTATAGGCATAATTACGAGGAAGCAGAACTTCTCCAGAAGCA
AGCGCACATGCGTTCCAAAATAAGAGCAAATTCGCTCTAAACACAGGAAAAGACCTGAAGCT
TTAATTAAGGGGTACATCCAACCCAGAGCGCTTTTGTGGGCACTGATTGCTCCAGCTTCT
GCGTCACTGCGCGAGGGAAGAGGGAAGAGGATCCAGGCGTTAGACATGTATAGACACAAAA
CAGCTGGAGATTGGGCTTAAAATACCCACCAAGCTCCAAAGAAGAGACCCAAGTCCCCAAA
CATTGATTT CAGGGCTGCCAGGAAGGAAGAGCAGCAGCAGGGTGGGAGAGAAGCTCCAGTCA
GCCCACAAGATGCCATTGTCCCCCGGCCTCCTGCTGCTGCTGCTCTCCGGGGCCACGGCCAC
CGCTGCCCTGCCCTGGAGGGTGGCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGG
CAGGAATAAGGAAAAGCAGCCTCCTGACTTTCCTCGCTTGGTGGTTTGAGTGACCTCCCAG
GCCAGTGCCGGGGCCCCTCATAGGAGAGGAAGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGC
ACCCCCCAGCAATCCGCGCGCCGGGACAGAATGCCCTGCAGGAAC'TCTTCTGGAAGACCT
TCTCCTCCTGCAAATAG

71/75

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA44175

><subunit 1 of 1, 155 aa, 1 stop

><MW: 17194, pI: 10.44, NX(S/T): 0

MYRHKNSWRLGLKYPPSSKEETQVPKTLISGLPGRKSSSRVGEKLQSAHKMPLSPGLLLLLL
SGATATAALPLEGGPTGRDSEHMQEAAGIRKSSLLTFLAWWFEWTSQASAGPLIGEEAREVA
RRQEGAPPQQSARRDRMPCRNFFWKTFSSCK

Important features of the protein:

Transmembrane domain:

amino acids 51-69

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 35-39, 92-96

N-myristoylation sites.

amino acids 64-70, 75-81, 90-96

Amidation site.

amino acids 33-37

FIGURE 71

GTCGTGTGCTTGGAGGAAGCCGCGGAACCCCCAGCGTCCGTCC**ATG**GCGTGGAGCCTTGGGA
 GCTGGCTGGGTGGCTGCCTGCTGGTGTGAGCATTGGGAATGGTACCACCTCCCGAAAATGTC
 AGAATGAATTCTGTTAATTTCAAGAACATTCTACAGTGGGAGTCACCTGCTTTTGCCAAAGG
 GAACCTGACTTTTACAGCTCAGTACCTAAGTTATAGGATATTCCAAGATAAATGCATGAATA
 CTACCTTGACGGAATGTGATTTCTCAAGTCTTTCCAAGTATGGTGACCACACCTTGAGAGTC
 AGGGCTGAATTTGCAGATGAGCATTGAGACTGGGTAAACATCACCTTCTGTCCTGTGGATGA
 CACCATTATTGGACCCCCCTGGAATGCAAGTAGAAGTACTTGCTGATTCTTTACATATGCGTT
 TCTTAGCCCCCTAAAATTGAGAATGAATACGAACTTGGACTATGAAGAATGTGTATAACTCA
 TGGACTTATAATGTGCAATACTGGAACAAACGGTACTGATGAAAAGTTTCAAATTACTCCCCA
 GTATGACTTTGAGGTCCTCAGAAACCTGGAGCCATGGACAACCTATTGTGTTCAAGTTCGAG
 GGTTTCTTCTGATCGGAACAAAGCTGGGGAATGGAGTGAGCCTGTCTGTGAGCAAACAACC
 CATGACGAAACGGTCCCCCTCCTGGATGGTGGCCGTCATCCTCATGGCCTCGGTCTTCATGGT
 CTGCCTGGCACTCCTCGGCTGCTTCTCCTTGCTGTGGTGCGTTTACAAGAAGACAAAGTACG
 CTTTCTCCCCTAGGAATTCTCTTCCACAGCACCTGAAAGAGTTTTTGGGCCATCCTCATCAT
 AACACACTTCTGTTTTTCTCCTTTCCATTGTCGGATGAGAATGATGTTTTTGACAAGCTAAG
 TGTCATTGCAGAAGACTCTGAGAGCGGCAAGCAGAATCCTGGTGACAGCTGCAGCCTCGGGA
 CCCCCTGGGCGAGGGGCCCAAGCT**TAG**GCTCTGAGAAGGAAACACACTCGGCTGGGCACA
 GTGACGTACTCCATCTCACATCTGCCTCAGTGAGGGATCAGGGCAGCAAACAAGGGCCAAGA
 CCATCTGAGCCAGCCCCACATCTAGAACTCCAGACCTGGACTTAGCCACCAGAGAGCTACAT
 TTAAAGGCTGTCTTGGCAAAAATACTCCATTTGGGAACTCACTGCCTTATAAAGGCTTTCA
 TGATGTTTTTCAAGAGTTGGCCACTGAGAGTGTAATTTTCAGCCTTTTATATCACTAAAATAA
 GATCATGTTTTAATTGTGAGAAACAGGGCCGAGCACAGTGGCTCACGCCTGTAATACCAGCA
 CCTTAGAGGTCGAGGCAGGCGGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAATA
 TGGTGAAACCCAGTCTCTACTAAAAATACAAAAATTAGCTAGGCATGATGGCGCATGCCTAT
 AATCCCAGCTACTCGAGTGCCTGAGGCAGGAGAATTGCATGAACCCGGGAGGAGGAGGAGGA
 GGTTGCAGTGAGCCGAGATAGCGGCACTGCACTCCAGCCTGGGTGACAAAGTGAGACTCCAT
 CTCAAAAAAAAAAAAAAAAAAATTGTGAGAAACAGAAATACTTAAATGAGGAATAAGAATGG
 AGATGTTACATCTGGTAGATGTAACATTCTACCAGATTATGGATGGACTGATCTGAAATCG
 ACCTCAACTCAAGGGTGGTCAGCTCAATGCTACACAGAGCACGGACTTTTGATTCTTTGCA
 GTACTTTGAATTTATTTTTCTACCTATATATGTTTTATATGCTGCTGGTGCTCCATTAAAGT
 TTTACTCTGTGTTGC

74/75

FIGURE 74

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA116510

><subunit 1 of 1, 494 aa, 1 stop

><MW: 54646, pI: 7.27, NX(S/T): 6

MARAAPLLAALTALLAAAAAGGDAPPGKIAVVGAGIGGSAVAHFLQQHFGPRVQIDVYEKGT
VGGRLATISV NKQHYESGAASFHSLSLHMQDFVKLLGLRHRREVVG RSAIFGGEHFMLEETD
WYLLNLFRLWWHYGISFLRLQMWVEEVM EKFMRIYKYQAHGYAFSGVEELLYSLGESTFVNM
TQHSVAESLLQVGVTQRFID DVVSAVLRASYGQSAAMPAFAGAMSLAGA QGSLWSVEGGNKL
VCSGLLKLTKANVIHATVTSVTLHSTEGKALYQVAYENEVGNSSDFYDIVVIATPLHLDNSS
SNLTFAGFHPPID DVQGSFQPTVVS LVHGYLNSSYFGFPDPKLF PFANILTTDFPSFFCTLD
NICPVNISASFRRKQPQEAAVWRVQSPKPLFRTQLKTLFRSYYSVQTA EWQAHP LYGSRPTL
PRFALHDQLFYLN ALEWAASSVEVM AAKNVALLAYNRWYQDL DKIDQKDLMHKVKTEL

Important features of the protein:**Signal peptide:**

amino acids 1-19

N-glycosylation sites.

amino acids 185-189, 290-294, 308-312, 312-316, 342-346, 378-382

N-myristoylation sites.

amino acids 33-39, 35-41, 38-44, 61-67, 64-70, 218-224, 234-240,
237-243, 429-435